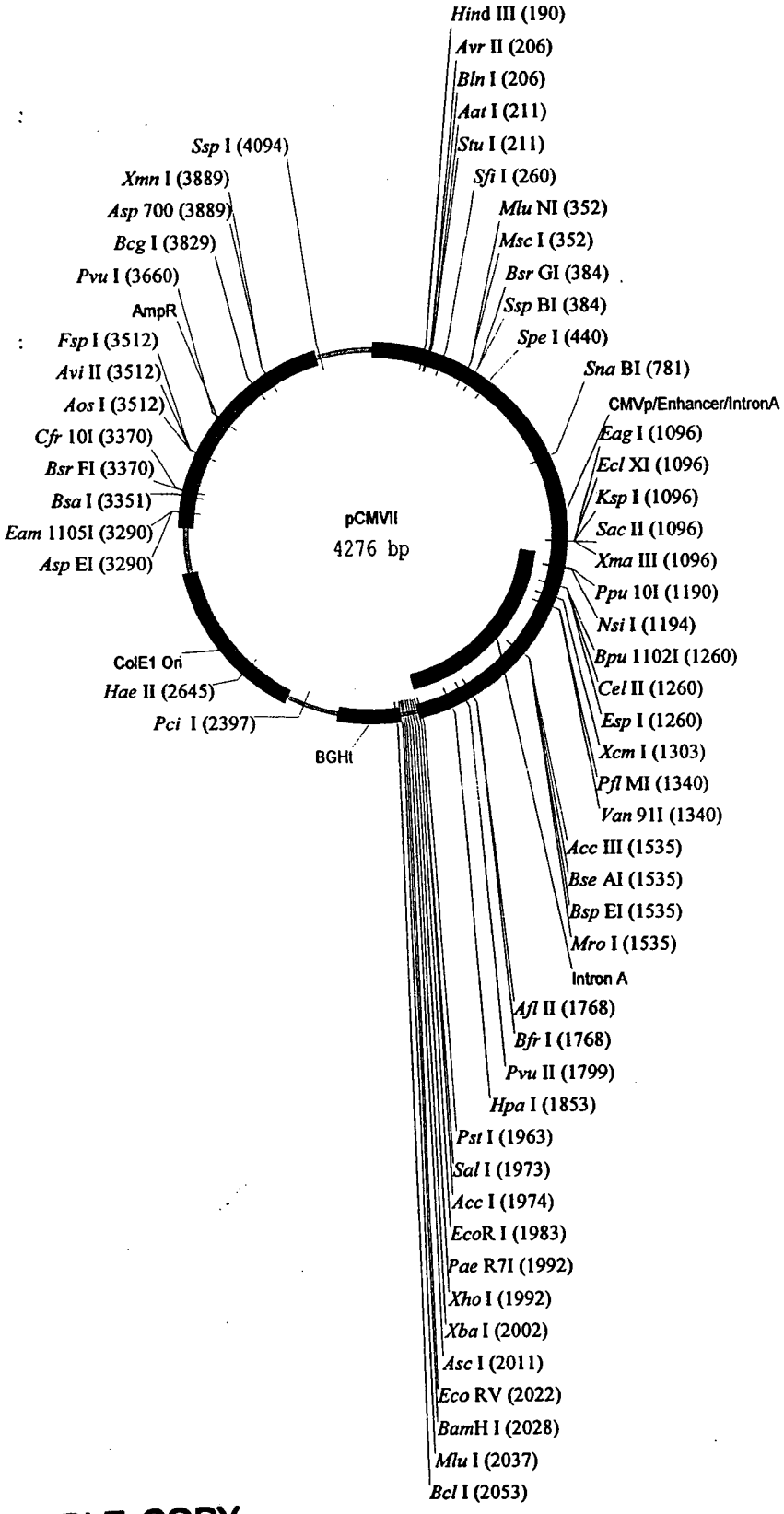

FIG. 1A

pCMV II

SEQ ID NO: 1

```
1   TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
    AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
    CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

101 TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
    AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                          HindIII
                                          --------
151 CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
    GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

201 AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG
    TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

251 AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
    TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301 TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
    ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351 GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
    CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401 CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
    GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451 TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG
    ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501 CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
    GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
    GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601 GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA
    CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651 CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
    GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701 TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA
    AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

751 CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
    GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801 CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG
    GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

851 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
    TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC
```

FIG. 1B

```
 901  TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC
      AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

 951  CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA
      GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001  GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC
      CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

1051  TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG
      ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101  GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA
      CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTCAT

1151  CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
      GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201  CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA
      GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

1251  TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC
      ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301  CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG
      GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351  CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT
      GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401  GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA
      CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT

1451  TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA
      AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

1501  TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT
      ATCGCACCCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551  CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651  CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
      GAATCCGTGT CGTGTTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701  GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
      CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

1751  GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
      CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

1801  GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG
      CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC
```

FIG. 1C

```
1851  TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
      AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901  CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
      GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

                          SalI        EcoRI      XhoI
                          ------      ------     ------
1951  GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCAGA CTCGAGCAAG
      CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTCT GAGCTCGTTC

       XbaI       AscI       EcoRV     BamHI      MluI
       ------   ----------   --------  --------   --------
2001  TCTAGAAAGG CGCGCCAAGA TATCAAGGAT CCACTACGCG TTAGAGCTCG
      AGATCTTTCC GCGCGGTTCT ATAGTTCCTA GGTGATGCGC AATCTCGAGC

2051  CTGATCAGCC TCGACTGTGC CTTCTAGTTG CCAGCCATCT GTTGTTTGCC
      GACTAGTCGG AGCTGACACG GAAGATCAAC GGTCGGTAGA CAACAAACGG

2101  CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG GTGCCACTCC CACTGTCCTT
      GGAGGGGGCA CGGAAGGAAC TGGGACCTTC CACGGTGAGG GTGACAGGAA

2151  TCCTAATAAA ATGAGGAAAT TGCATCGCAT TGTCTGAGTA GGTGTCATTC
      AGGATTATTT TACTCCTTTA ACGTAGCGTA ACAGACTCAT CCACAGTAAG

2201  TATTCTGGGG GGTGGGGTGG GGCAGGACAG CAAGGGGGAG GATTGGGAAG
      ATAAGACCCC CCACCCCACC CCGTCCTGTC GTTCCCCCTC CTAACCCTTC

2251  ACAATAGCAG GCATGCTGGG GAGCTCTTCC GCTTCCTCGC TCACTGACTC
      TGTTATCGTC CGTACGACCC CTCGAGAAGG CGAAGGAGCG AGTGACTGAG

2301  GCTGCGCTCG GTCGTTCGGC TGCGGCGAGC GGTATCAGCT CACTCAAAGG
      CGACGCGAGC CAGCAAGCCG ACGCCGCTCG CCATAGTCGA GTGAGTTTCC

2351  CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG AAAGAACATG
      GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC TTTCTTGTAC

2401  TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG CCGCGTTGCT
      ACTCGTTTTC CGGTCGTTTT CCGGTCCTTG GCATTTTTCC GGCGCAACGA

2451  GGCGTTTTTC CATAGGCTCC GCCCCCCTGA CGAGCATCAC AAAAATCGAC
      CCGCAAAAAG GTATCCGAGG CGGGGGGACT GCTCGTAGTG TTTTTAGCTG

2501  GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG
      CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATATTTC TATGGTCCGC

2551  TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA CCCTGCCGCT
      AAAGGGGGAC CTTCGAGGGA GCACGCGAGA GGACAAGGCT GGGACGGCGA

2601  TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG GCGCTTTCTC
      ATGGCCTATG GACAGGCGGA AAGAGGGAAG CCCTTCGCAC CGCGAAAGAG

2651  AATGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT TCGCTCCAAG
      TTACGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA AGCGAGGTTC

2701  CTGGGCTGTG TGCACGAACC CCCCGTTCAG CCCGACCGCT GCGCCTTATC
      GACCCGACAC ACGTGCTTGG GGGGCAAGTC GGGCTGGCGA CGCGGAATAG
```

FIG. 1D

```
2751  CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC TTATCGCCAC
      GCCATTGATA GCAGAACTCA GGTTGGGCCA TTCTGTGCTG AATAGCGGTG

2801  TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA TGTAGGCGGT
      ACCGTCGTCG GTGACCATTG TCCTAATCGT CTCGCTCCAT ACATCCGCCA

2851  GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGGAC
      CGATGTCTCA AGAACTTCAC CACCGGATTG ATGCCGATGT GATCTTCCTG

2901  AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG
      TCATAAACCA TAGACGCGAG ACGACTTCGG TCAATGGAAG CCTTTTTCTC

2951  TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG CGGTGGTTTT
      AACCATCGAG AACTAGGCCG TTTGTTTGGT GGCGACCATC GCCACCAAAA

3001  TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAAGAAGA
      AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA GAGTTCTTCT

3051  TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC
      AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG

3101  GTTAAGGGAT TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC
      CAATTCCCTA AAACCAGTAC TCTAATAGTT TTTCCTAGAA GTGGATCTAG

3151  CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
      GAAAATTTAA TTTTTACTTC AAAATTTAGT TAGATTTCAT ATATACTCAT

3201  AACTTGGTCT GACAGTTACC AATGCTTAAT CAGTGAGGCA CCTATCTCAG
      TTGAACCAGA CTGTCAATGG TTACGAATTA GTCACTCCGT GGATAGAGTC

3251  CGATCTGTCT ATTTCGTTCA TCCATAGTTG CCTGACTCCC CGTCGTGTAG
      GCTAGACAGA TAAAGCAAGT AGGTATCAAC GGACTGAGGG GCAGCACATC

3301  ATAACTACGA TACGGGAGGG CTTACCATCT GGCCCCAGTG CTGCAATGAT
      TATTGATGCT ATGCCCTCCC GAATGGTAGA CCGGGGTCAC GACGTTACTA

3351  ACCGCGAGAC CCACGCTCAC CGGCTCCAGA TTTATCAGCA ATAAACCAGC
      TGGCGCTCTG GGTGCGAGTG GCCGAGGTCT AAATAGTCGT TATTTGGTCG

3401  CAGCCGGAAG GGCCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCCTCC
      GTCGGCCTTC CCGGCTCGCG TCTTCACCAG GACGTTGAAA TAGGCGGAGG

3451  ATCCAGTCTA TTAATTGTTG CCGGGAAGCT AGAGTAAGTA GTTCGCCAGT
      TAGGTCAGAT AATTAACAAC GGCCCTTCGA TCTCATTCAT CAAGCGGTCA

3501  TAATAGTTTG CGCAACGTTG TTGCCATTGC TACAGGCATC GTGGTGTCAC
      ATTATCAAAC GCGTTGCAAC AACGGTAACG ATGTCCGTAG CACCACAGTG

3551  GCTCGTCGTT TGGTATGGCT TCATTCAGCT CCGGTTCCCA ACGATCAAGG
      CGAGCAGCAA ACCATACCGA AGTAAGTCGA GGCCAAGGGT TGCTAGTTCC

3601  CGAGTTACAT GATCCCCCAT GTTGTGCAAA AAAGCGGTTA GCTCCTTCGG
      GCTCAATGTA CTAGGGGGTA CAACACGTTT TTTCGCCAAT CGAGGAAGCC

3651  TCCTCCGATC GTTGTCAGAA GTAAGTTGGC CGCAGTGTTA TCACTCATGG
      AGGAGGCTAG CAACAGTCTT CATTCAACCG GCGTCACAAT AGTGAGTACC
```

FIG. 1E

```
3701  TTATGGCAGC ACTGCATAAT TCTCTTACTG TCATGCCATC CGTAAGATGC
      AATACCGTCG TGACGTATTA AGAGAATGAC AGTACGGTAG GCATTCTACG

3751  TTTTCTGTGA CTGGTGAGTA CTCAACCAAG TCATTCTGAG AATAGTGTAT
      AAAAGACACT GACCACTCAT GAGTTGGTTC AGTAAGACTC TTATCACATA

3801  GCGGCGACCG AGTTGCTCTT GCCCGGCGTC AATACGGGAT AATACCGCGC
      CGCCGCTGGC TCAACGAGAA CGGGCCGCAG TTATGCCCTA TTATGGCGCG

3851  CACATAGCAG AACTTTAAAA GTGCTCATCA TTGGAAAACG TTCTTCGGGG
      GTGTATCGTC TTGAAATTTT CACGAGTAGT AACCTTTTGC AAGAAGCCCC

3901  CGAAAACTCT CAAGGATCTT ACCGCTGTTG AGATCCAGTT CGATGTAACC
      GCTTTTGAGA GTTCCTAGAA TGGCGACAAC TCTAGGTCAA GCTACATTGG

3951  CACTCGTGCA CCCAACTGAT CTTCAGCATC TTTTACTTTC ACCAGCGTTT
      GTGAGCACGT GGGTTGACTA GAAGTCGTAG AAAATGAAAG TGGTCGCAAA

4001  CTGGGTGAGC AAAAACAGGA AGGCAAAATG CCGCAAAAAA GGGAATAAGG
      GACCCACTCG TTTTTGTCCT TCCGTTTTAC GGCGTTTTTT CCCTTATTCC

4051  GCGACACGGA AATGTTGAAT ACTCATACTC TTCCTTTTTC AATATTATTG
      CGCTGTGCCT TTACAACTTA TGAGTATGAG AAGGAAAAAG TTATAATAAC

4101  AAGCATTTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA
      TTCGTAAATA GTCCCAATAA CAGAGTACTC GCCTATGTAT AAACTTACAT

4151  TTTAGAAAAA TAAACAAATA GGGGTTCCGC GCACATTTCC CCGAAAAGTG
      AAATCTTTTT ATTTGTTTAT CCCCAAGGCG CGTGTAAAGG GGCTTTTCAC

4201  CCACCTGACG TCTAAGAAAC CATTATTATC ATGACATTAA CCTATAAAAA
      GGTGGACTGC AGATTCTTTG GTAATAATAG TACTGTAATT GGATATTTTT

4251  TAGGCGTATC ACGAGGCCCT TTCGTC
      ATCCGCATAG TGCTCCGGGA AAGCAG
```

FIG. 1F

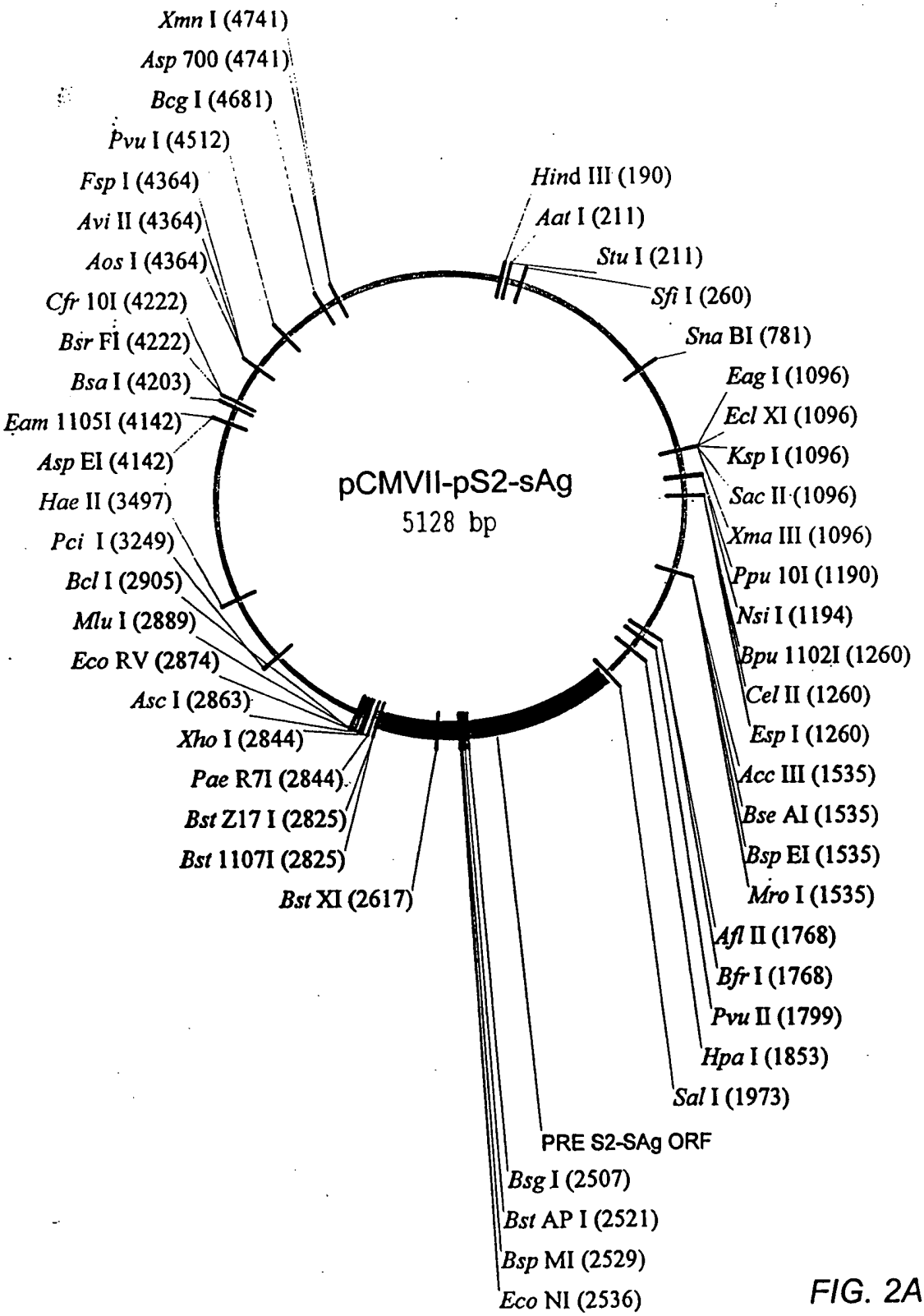
pCMVII-pS2-sAg
5128 bp
Xmn I (4741)
Asp 700 (4741)
Bcg I (4681)
Pvu I (4512)
Fsp I (4364)
Avi II (4364)
Aos I (4364)
Cfr 10I (4222)
Bsr FI (4222)
Bsa I (4203)
Eam 1105I (4142)
Asp EI (4142)
Hae II (3497)
Pci I (3249)
Bcl I (2905)
Mlu I (2889)
Eco RV (2874)
Asc I (2863)
Xho I (2844)
Pae R7I (2844)
Bst Z17 I (2825)
Bst 1107I (2825)
Bst XI (2617)
Hind III (190)
Aat I (211)
Stu I (211)
Sfi I (260)
Sna BI (781)
Eag I (1096)
Ecl XI (1096)
Ksp I (1096)
Sac II (1096)
Xma III (1096)
Ppu 10I (1190)
Nsi I (1194)
Bpu 1102I (1260)
Cel II (1260)
Esp I (1260)
Acc III (1535)
Bse AI (1535)
Bsp EI (1535)
Mro I (1535)
Afl II (1768)
Bfr I (1768)
Pvu II (1799)
Hpa I (1853)
Sal I (1973)
PRE S2-SAg ORF
Bsg I (2507)
Bst AP I (2521)
Bsp MI (2529)
Eco NI (2536)

FIG. 2A

SEQ ID NO:2

```
1    TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
     AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

51   GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
     CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
     AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                              HindIII
                                              ~~~~~~~
151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
     GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

           StuI
           ~~~~~~~
           AatI
           ~~~~~~~
201  AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG
     TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

         SfiI
         ~~~~~~~~~~~~~~~
251  AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
     TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301  TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
     ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351  GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
     CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
     GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451  TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG
     ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501  CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
     GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551  CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
     GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601  GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA
     CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651  CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
     GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA
     AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

                                  SnaBI
                                  ~~~~~~~
751  CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
     GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG
```

FIG. 2B

```
801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG
     GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

851  ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
     TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

901  TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC
     AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

951  CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA
     GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001 GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC
     CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

                                                  XmaIII
                                                  ~~~~~~
                                                SacII
                                               ~~~~~~
                                               KspI
                                               ~~~~~~
                                                   EclXI
                                                   ~~~~~~
                                                   EagI
                                                   ~~~~~~
1051 TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG
     ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101 GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA
     CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTCAT

                                                Ppu10I
                                                ~~~~~~~
                                                NsiI
                                                ~~~~~~~
1151 CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
     GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA
     GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

              EspI
              ~~~~~~~~
              CelII
              ~~~~~~~~
            Bpu1102I
            ~~~~~~~~
1251 TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC
     ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301 CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG
     GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351 CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT
     GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401 GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA
     CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT
```

FIG. 2C

```
1451  TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA
      AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

                                          MroI
                                          ~~~~~~
                                          BspEI
                                          ~~~~~~
                                          BseAI
                                          ~~~~~~
                                          AccIII
                                          ~~~~~~
1501  TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT
      ATCGCACCCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551  CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651  CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
      GAATCCGTGT CGTGTTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701  GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
      CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

                      BfrI
                      ~~~~~~~~
                      AflII                          PvuII
                      ~~~~~~~                        ~~~~~
1751  GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
      CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

      PvuII                                          HpaI
      ~                                              ~
1801  GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG
      CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

      HpaI
      ~~~~~
1851  TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
      AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901  CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
      GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

  +2                                  SEQ ID NO:3 M  Q  W  N
                                  SalI
                                  ~~~~~~
1951  GTCTTTTCTG CAGTCACCGT CGTCGACCTA AGAATTCATG CAGTGGAACT
      CAGAAAAGAC GTCAGTGGCA GCAGCTGGAT TCTTAAGTAC GTCACCTTGA

  +2  S  T  A  F  H  Q  T  L  Q  D  P  R  V  R  G  L  Y
2001  CCACTGCCTT CCACCAAACT CTGCAGGATC CCAGAGTCAG GGGTCTGTAT
      GGTGACGGAA GGTGGTTTGA GACGTCCTAG GGTCTCAGTC CCCAGACATA

  +2  L  P  A  G  G  S  S  S  G  T  V  N  P  A  P  N  I
2051  CTTCCTGCTG GTGGCTCCAG TTCAGGAACA GTAAACCCTG CTCCGAATAT
      GAAGGACGAC CACCGAGGTC AAGTCCTTGT CATTTGGGAC GAGGCTTATA
```

FIG. 2D

```
  +2   A  S  H  I  S  S  I  S  A  R  T  G  D  P  V  T
2101  TGCCTCTCAC ATCTCGTCAA TCTCCGCGAG GACTGGGGAC CCTGTGACGA
      ACGGAGAGTG TAGAGCAGTT AGAGGCGCTC CTGACCCCTG GGACACTGCT

  +2 N  M  E  N  I  T  S  G  F  L  G  P  L  L  V  L  Q
2151  ACATGGAGAA CATCACATCA GGATTCCTAG GACCCCTGCT CGTGTTACAG
      TGTACCTCTT GTAGTGTAGT CCTAAGGATC CTGGGGACGA GCACAATGTC

  +2  A  G  F  F  L  L  T  R  I  L  T  I  P  Q  S  L  D
2201  GCGGGGTTTT TCTTGTTGAC AAGAATCCTC ACAATACCGC AGAGTCTAGA
      CGCCCCAAAA AGAACAACTG TTCTTAGGAG TGTTATGGCG TCTCAGATCT

  +2  S  W  W  T  S  L  N  F  L  G  G  S  P  V  C  L
2251  CTCGTGGTGG ACTTCTCTCA ATTTTCTAGG GGGATCTCCC GTGTGTCTTG
      GAGCACCACC TGAAGAGAGT TAAAAGATCC CCCTAGAGGG CACACAGAAC

  +2 G  Q  N  S  Q  S  P  T  S  N  H  S  P  T  S  C  P
2301  GCCAAAATTC GCAGTCCCCA ACCTCCAATC ACTCACCAAC CTCCTGTCCT
      CGGTTTTAAG CGTCAGGGGT TGGAGGTTAG TGAGTGGTTG GAGGACAGGA

  +2  P  I  C  P  G  Y  R  W  M  C  L  R  R  F  I  I  F
2351  CCAATTTGTC CTGGTTATCG CTGGATGTGT CTGCGGCGTT TTATCATATT
      GGTTAAACAG GACCAATAGC GACCTACACA GACGCCGCAA AATAGTATAA

  +2  L  F  I  L  L  L  C  L  I  F  L  L  V  L  L  D
2401  CCTCTTCATC CTGCTGCTAT GCCTCATCTT CTTATTGGTT CTTCTGGATT
      GGAGAAGTAG GACGACGATA CGGAGTAGAA GAATAACCAA GAAGACCTAA

  +2 Y  Q  G  M  L  P  V  C  P  L  I  P  G  S  T  T  T
2451  ATCAAGGTAT GTTGCCCGTT TGTCCTCTAA TTCCAGGATC AACAACAACC
      TAGTTCCATA CAACGGGCAA ACAGGAGATT AAGGTCCTAG TTGTTGTTGG

  +2  S  T  G  P  C  K  T  C  T  T  P  A  Q  G  N  S  M
                    BstAP  I
                   ~~~~~~~~~~~~
                         BspMI          EcoNI
                        ~~~~~~~        ~~~~~~~~~~~~
2501  AGTACGGGAC CATGCAAAAC CTGCACGACT CCTGCTCAAG GCAACTCTAT
      TCATGCCCTG GTACGTTTTG GACGTGCTGA GGACGAGTTC CGTTGAGATA
                              BsgI
                             ~~~~~~

  +2   F  P  S  C  C  C  T  K  P  T  D  G  N  C  T  C
2551  GTTTCCCTCA TGTTGCTGTA CAAAACCTAC GGATGGAAAT TGCACCTGTA
      CAAAGGGAGT ACAACGACAT GTTTTGGATG CCTACCTTTA ACGTGGACAT

  +2 I  P  I  P  S  S  W  A  F  A  K  Y  L  W  E  W  A
                   BstXI
                ~~~~~~~~~~~~~~
2601  TTCCCATCCC ATCGTCCTGG GCTTTCGCAA AATACCTATG GGAGTGGGCC
      AAGGGTAGGG TAGCAGGACC CGAAAGCGTT TTATGGATAC CCTCACCCGG

  +2  S  V  R  F  S  W  L  S  L  L  V  P  F  V  Q  W  F
2651  TCAGTCCGTT TCTCTTGGCT CAGTTTACTA GTGCCATTTG TTCAGTGGTT
      AGTCAGGCAA AGAGAACCGA GTCAAATGAT CACGGTAAAC AAGTCACCAA

  +2   V  G  L  S  P  T  V  W  L  S  A  I  W  M  M  W
2701  CGTAGGGCTT TCCCCCACTG TTTGGCTTTC AGCTATATGG ATGATGTGGT
      GCATCCCGAA AGGGGGTGAC AAACCGAAAG TCGATATACC TACTACACCA
```

FIG. 2E

```
   +2 Y  W  G  P   S  L  Y   S  I  V  S   P  F  I   P  L  L
2751  ATTGGGGGCC AAGTCTGTAC AGCATCGTGA GTCCCTTTAT ACCGCTGTTA
      TAACCCCCGG TTCAGACATG TCGTAGCACT CAGGGAAATA TGGCGACAAT

   +2  P  I  F  F   C  L  W   V  Y  I   *
                             BstZ17 I                 XhoI
                              ~~~~~~                  ~~~~~~
                             Bst1107I                 PaeR7I
                              ~~~~~~                  ~~~~~~
2801  CCAATTTTCT TTTGTCTCTG GGTATACATT TAAGAATTCA GACTCGAGCA
      GGTTAAAAGA AAACAGAGAC CCATATGTAA ATTCTTAAGT CTGAGCTCGT

                    AscI       EcoRV              MluI
                   ~~~~~~~~   ~~~~~~             ~~~~~~~
2851  AGTCTAGAAA GGCGCGCCAA GATATCAAGG ATCCACTACG CGTTAGAGCT
      TCAGATCTTT CCGCGCGGTT CTATAGTTCC TAGGTGATGC GCAATCTCGA

          BclI
         ~~~~~~
2901  CGCTGATCAG CCTCGACTGT GCCTTCTAGT TGCCAGCCAT CTGTTGTTTG
      GCGACTAGTC GGAGCTGACA CGGAAGATCA ACGGTCGGTA GACAACAAAC

2951  CCCCTCCCCC GTGCCTTCCT TGACCCTGGA AGGTGCCACT CCCACTGTCC
      GGGGAGGGGG CACGGAAGGA ACTGGGACCT TCCACGGTGA GGGTGACAGG

3001  TTTCCTAATA AAATGAGGAA ATTGCATCGC ATTGTCTGAG TAGGTGTCAT
      AAAGGATTAT TTTACTCCTT TAACGTAGCG TAACAGACTC ATCCACAGTA

3051  TCTATTCTGG GGGGTGGGGT GGGGCAGGAC AGCAAGGGGG AGGATTGGGA
      AGATAAGACC CCCCACCCCA CCCCGTCCTG TCGTTCCCCC TCCTAACCCT

3101  AGACAATAGC AGGCATGCTG GGGAGCTCTT CCGCTTCCTC GCTCACTGAC
      TCTGTTATCG TCCGTACGAC CCCTCGAGAA GGCGAAGGAG CGAGTGACTG

3151  TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA GCGGTATCAG CTCACTCAAA
      AGCGACGCGA GCCAGCAAGC CGACGCCGCT CGCCATAGTC GAGTGAGTTT

                                                      Pci I
                                                       ~~~
3201  GGCGGTAATA CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA
      CCGCCATTAT GCCAATAGGT GTCTTAGTCC CCTATTGCGT CCTTTCTTGT

      Pci I
      ~~~
3251  TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG
      ACACTCGTTT TCCGGTCGTT TTCCGGTCCT TGGCATTTTT CCGGCGCAAC

3301  CTGGCGTTTT TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG
      GACCGCAAAA AGGTATCCGA GGCGGGGGGA CTGCTCGTAG TGTTTTTAGC

3351  ACGCTCAAGT CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG
      TGCGAGTTCA GTCTCCACCG CTTTGGGCTG TCCTGATATT TCTATGGTCC

3401  CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG
      GCAAAGGGGG ACCTTCGAGG GAGCACGCGA GAGGACAAGG CTGGGACGGC
```

FIG. 2F

```
                                                      HaeII
                                                      ~~~~~~
3451  CTTACCGGAT ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC
      GAATGGCCTA TGGACAGGCG GAAAGAGGGA AGCCCTTCGC ACCGCGAAAG

3501  TCAATGCTCA CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA
      AGTTACGAGT GCGACATCCA TAGAGTCAAG CCACATCCAG CAAGCGAGGT

3551  AGCTGGGCTG TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA
      TCGACCCGAC ACACGTGCTT GGGGGGCAAG TCGGGCTGGC GACGCGGAAT

3601  TCCGGTAACT ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC
      AGGCCATTGA TAGCAGAACT CAGGTTGGGC CATTCTGTGC TGAATAGCGG

3651  ACTGGCAGCA GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG
      TGACCGTCGT CGGTGACCAT TGTCCTAATC GTCTCGCTCC ATACATCCGC

3701  GTGCTACAGA GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGG
      CACGATGTCT CAAGAACTTC ACCACCGGAT TGATGCCGAT GTGATCTTCC

3751  ACAGTATTTG GTATCTGCGC TCTGCTGAAG CCAGTTACCT TCGGAAAAAG
      TGTCATAAAC CATAGACGCG AGACGACTTC GGTCAATGGA AGCCTTTTTC

3801  AGTTGGTAGC TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT
      TCAACCATCG AGAACTAGGC CGTTTGTTTG GTGGCGACCA TCGCCACCAA

3851  TTTTTGTTTG CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA
      AAAAACAAAC GTTCGTCGTC TAATGCGCGT CTTTTTTTCC TAGAGTTCTT

3901  GATCCTTTGA TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC
      CTAGGAAACT AGAAAAGATG CCCCAGACTG CGAGTCACCT TGCTTTTGAG

3951  ACGTTAAGGG ATTTTGGTCA TGAGATTATC AAAAAGGATC TTCACCTAGA
      TGCAATTCCC TAAAACCAGT ACTCTAATAG TTTTTCCTAG AAGTGGATCT

4001  TCCTTTTAAA TTAAAAATGA AGTTTTAAAT CAATCTAAAG TATATATGAG
      AGGAAAATTT AATTTTTACT TCAAAATTTA GTTAGATTTC ATATATACTC

4051  TAAACTTGGT CTGACAGTTA CCAATGCTTA ATCAGTGAGG CACCTATCTC
      ATTTGAACCA GACTGTCAAT GGTTACGAAT TAGTCACTCC GTGGATAGAG

                                          Eam1105I
                                          ~~~~~~~~~~~~
                                            AspEI
                                            ~~~~~~~~~~~~
4101  AGCGATCTGT CTATTTCGTT CATCCATAGT TGCCTGACTC CCCGTCGTGT
      TCGCTAGACA GATAAAGCAA GTAGGTATCA ACGGACTGAG GGGCAGCACA

4151  AGATAACTAC GATACGGGAG GGCTTACCAT CTGGCCCCAG TGCTGCAATG
      TCTATTGATG CTATGCCCTC CCGAATGGTA GACCGGGGTC ACGACGTTAC
```

FIG. 2G

```
                          Cfr10I
                          ~~~~~~
                           BsrFI
                          ~~~~~~
4201  ATACCGCGAG ACCCACGCTC ACCGGCTCCA GATTTATCAG CAATAAACCA
      TATGGCGCTC TGGGTGCGAG TGGCCGAGGT CTAAATAGTC GTTATTTGGT
             BsaI
            ~~~~~~~
--------------------------------------------------------------

4251  GCCAGCCGGA AGGGCCGAGC GCAGAAGTGG TCCTGCAACT TTATCCGCCT
      CGGTCGGCCT TCCCGGCTCG CGTCTTCACC AGGACGTTGA AATAGGCGGA
--------------------------------------------------------------

4301  CCATCCAGTC TATTAATTGT TGCCGGGAAG CTAGAGTAAG TAGTTCGCCA
      GGTAGGTCAG ATAATTAACA ACGGCCCTTC GATCTCATTC ATCAAGCGGT
--------------------------------------------------------------

                  FspI
                 ~~~~~~
                  AviII
                 ~~~~~~
                  AosI
                 ~~~~~~
4351  GTTAATAGTT TGCGCAACGT TGTTGCCATT GCTACAGGCA TCGTGGTGTC
      CAATTATCAA ACGCGTTGCA ACAACGGTAA CGATGTCCGT AGCACCACAG
--------------------------------------------------------------

4401  ACGCTCGTCG TTTGGTATGG CTTCATTCAG CTCCGGTTCC CAACGATCAA
      TGCGAGCAGC AAACCATACC GAAGTAAGTC GAGGCCAAGG GTTGCTAGTT
--------------------------------------------------------------

4451  GGCGAGTTAC ATGATCCCCC ATGTTGTGCA AAAAAGCGGT TAGCTCCTTC
      CCGCTCAATG TACTAGGGGG TACAACACGT TTTTTCGCCA ATCGAGGAAG
--------------------------------------------------------------

               PvuI
              ~~~~~~~
4501  GGTCCTCCGA TCGTTGTCAG AAGTAAGTTG GCCGCAGTGT TATCACTCAT
      CCAGGAGGCT AGCAACAGTC TTCATTCAAC CGGCGTCACA ATAGTGAGTA
--------------------------------------------------------------

4551  GGTTATGGCA GCACTGCATA ATTCTCTTAC TGTCATGCCA TCCGTAAGAT
      CCAATACCGT CGTGACGTAT TAAGAGAATG ACAGTACGGT AGGCATTCTA
--------------------------------------------------------------

4601  GCTTTTCTGT GACTGGTGAG TACTCAACCA AGTCATTCTG AGAATAGTGT
      CGAAAAGACA CTGACCACTC ATGAGTTGGT TCAGTAAGAC TCTTATCACA
--------------------------------------------------------------

                .BcgI
              ~~~~~~~~~~~~~~
4651  ATGCGGCGAC CGAGTTGCTC TTGCCCGGCG TCAATACGGG ATAATACCGC
      TACGCCGCTG GCTCAACGAG AACGGGCCGC AGTTATGCCC TATTATGGCG
--------------------------------------------------------------

                                          XmnI
                                         ~~~~~~~~~~~~
                                          Asp700
                                         ~~~~~~~~~~~~
4701  GCCACATAGC AGAACTTTAA AAGTGCTCAT CATTGGAAAA CGTTCTTCGG
      CGGTGTATCG TCTTGAAATT TTCACGAGTA GTAACCTTTT GCAAGAAGCC
--------------------------------------------------------------

4751  GGCGAAAACT CTCAAGGATC TTACCGCTGT TGAGATCCAG TTCGATGTAA
      CCGCTTTTGA GAGTTCCTAG AATGGCGACA ACTCTAGGTC AAGCTACATT
--------------------------------------------------------------

4801  CCCACTCGTG CACCCAACTG ATCTTCAGCA TCTTTTACTT TCACCAGCGT
      GGGTGAGCAC GTGGGTTGAC TAGAAGTCGT AGAAAATGAA AGTGGTCGCA
--------------------------------------------------------------
```

FIG. 2H

```
4851  TTCTGGGTGA GCAAAAACAG GAAGGCAAAA TGCCGCAAAA AAGGGAATAA
      AAGACCCACT CGTTTTTGTC CTTCCGTTTT ACGGCGTTTT TTCCCTTATT

4901  GGGCGACACG GAAATGTTGA ATACTCATAC TCTTCCTTTT TCAATATTAT
      CCCGCTGTGC CTTTACAACT TATGAGTATG AGAAGGAAAA AGTTATAATA

4951  TGAAGCATTT ATCAGGGTTA TTGTCTCATG AGCGGATACA TATTTGAATG
      ACTTCGTAAA TAGTCCCAAT AACAGAGTAC TCGCCTATGT ATAAACTTAC

5001  TATTTAGAAA AATAAACAAA TAGGGGTTCC GCGCACATTT CCCCGAAAAG
      ATAAATCTTT TTATTTGTTT ATCCCCAAGG CGCGTGTAAA GGGGCTTTTC

5051  TGCCACCTGA CGTCTAAGAA ACCATTATTA TCATGACATT AACCTATAAA
      ACGGTGGACT GCAGATTCTT TGGTAATAAT AGTACTGTAA TTGGATATTT

5101  AATAGGCGTA TCACGAGGCC CTTTCGTC
      TTATCCGCAT AGTGCTCCGG GAAAGCAG
```

FIG. 2I

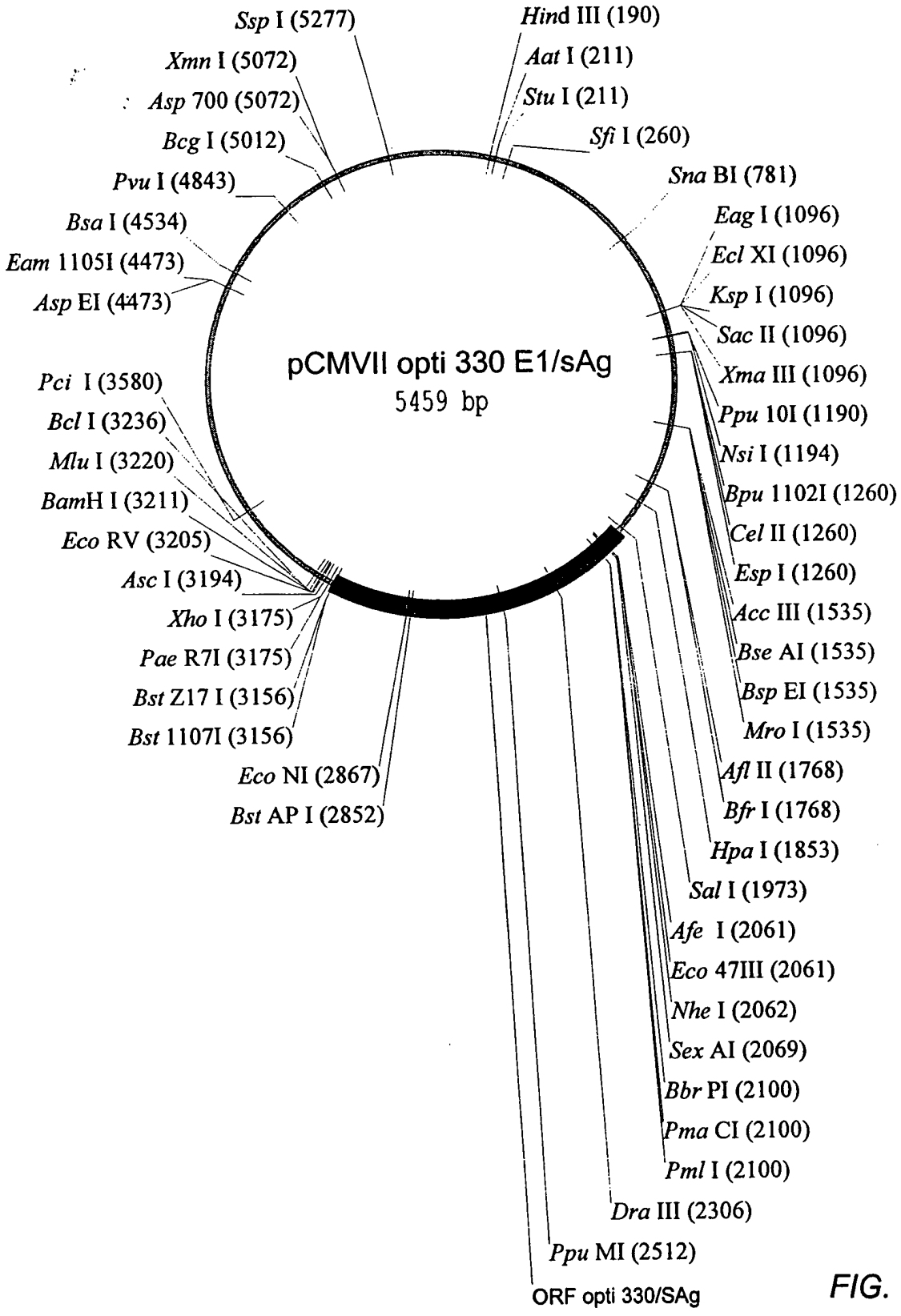
pCMVII opti 330 E1/sAg
5459 bp
Ssp I (5277)
Xmn I (5072)
Asp 700 (5072)
Bcg I (5012)
Pvu I (4843)
Bsa I (4534)
Eam 1105I (4473)
Asp EI (4473)
Pci I (3580)
Bcl I (3236)
Mlu I (3220)
BamH I (3211)
Eco RV (3205)
Asc I (3194)
Xho I (3175)
Pae R7I (3175)
Bst Z17 I (3156)
Bst 1107I (3156)
Eco NI (2867)
Bst AP I (2852)
Hind III (190)
Aat I (211)
Stu I (211)
Sfi I (260)
Sna BI (781)
Eag I (1096)
Ecl XI (1096)
Ksp I (1096)
Sac II (1096)
Xma III (1096)
Ppu 10I (1190)
Nsi I (1194)
Bpu 1102I (1260)
Cel II (1260)
Esp I (1260)
Acc III (1535)
Bse AI (1535)
Bsp EI (1535)
Mro I (1535)
Afl II (1768)
Bfr I (1768)
Hpa I (1853)
Sal I (1973)
Afe I (2061)
Eco 47III (2061)
Nhe I (2062)
Sex AI (2069)
Bbr PI (2100)
Pma CI (2100)
Pml I (2100)
Dra III (2306)
Ppu MI (2512)
ORF opti 330/SAg

FIG. 3A

SEQ ID No: 4

```
1    TCGCGCGTTT CGGTGATGAC  TGAAAACC TCTGACACAT GCAGCTCCCG
     AGCGCGCAAA GCCACTAC    ACTTTTGG AGACTGTGTA CGTCGAGGGC

51   GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
     CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
     AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                              HindIII
                                              ~~~~~~~
151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA
     GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT

           StuI
           ~~~~~~~
           AatI
           ~~~~~~~
201  AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG AATAGCTCAG
     TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC TTATCGAGTC

         SfiI
         ~~~~~~~~~~~~~~~
251  AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA
     TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT

301  TGGGGCGGAG AATGGGCGGA ACTGGGCGGG GAGGGAATTA TTGGCTATTG
     ACCCCGCCTC TTACCCGCCT TGACCCGCCC CTCCCTTAAT AACCGATAAC

351  GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
     CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA
     GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT

451  TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA TGGAGTTCCG
     ATCATTAGTT AATGCCCCAG TAATCAAGTA TCGGGTATAT ACCTCAAGGC

501  CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
     GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG

551  CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA
     GGGCGGGTAA CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT

601  GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA
     CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA TTTGACGGGT

651  CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG
     GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC

701  TCAATGACGG TAAATGGCCC GCCTGGCATT ATGCCCAGTA CATGACCTTA
     AGTTACTGCC ATTTACCGGG CGGACCGTAA TACGGGTCAT GTACTGGAAT

                                   SnaBI
                                   ~~~~~~~
751  CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
     GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG
```

FIG. 3B

```
 801  CATGGTGATG CGGTTTT  AGTACACCAA TGGGCGTGGA TAGCGGTTTG
      GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC

 851  ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG
      TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT AACTGCAGTT ACCCTCAAAC

 901  TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC
      AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG

 951  CCCGTTGACG CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA
      GGGCAACTGC GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT

1001  GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG CCATCCACGC
      CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC GGTAGGTGCG

                                                    XmaIII
                                                    ~~~~~~
                                                  SacII
                                                  ~~~~~~
                                                  KspI
                                                  ~~~~~~
                                                      EclXI
                                                      ~~~~~~
                                                      EagI
                                                      ~~~~~~
1051  TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG
      ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC

1101  GGAACGGTGC ATTGGAACGC GGATTCCCCG TGCCAAGAGT GACGTAAGTA
      CCTTGCCACG TAACCTTGCG CCTAAGGGGC ACGGTTCTCA CTGCATTCAT

                                                Ppu10I
                                                ~~~~~~~
                                                 NsiI
                                                 ~~~~~~~
1151  CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
      GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201  CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA
      GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT

            EspI
            ~~~~~~~~
            CelII
            ~~~~~~~~
          Bpu1102I
          ~~~~~~~~
1251  TGGTATAGCT TAGCCTATAG GTGTGGGTTA TTGACCATTA TTGACCACTC
      ACCATATCGA ATCGGATATC CACACCCAAT AACTGGTAAT AACTGGTGAG

1301  CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG
      GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC

1351  CCACAACTAT CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT
      GGTGTTGATA GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA

1401  GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT TATTTACAAA
      CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA ATAAATGTTT
```

FIG. 3C

```
1451  TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA
      AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT

                                          MroI
                                          ~~~~~~
                                          BspEI
                                          ~~~~~~
                                          BseAI
                                          ~~~~~~
                                          AccIII
                                          ~~~~~~
1501  TAGCGTGGGA TCTCCGACAT CTCGGGTACG TGTTCCGGAC ATGGGCTCTT
      ATCGCACCCT AGAGGCTGTA GAGCCCATGC ACAAGGCCTG TACCCGAGAA

1551  CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT

1651  CTTAGGCACA GCACAATGCC CACCACCACC AGTGTGCCGC ACAAGGCCGT
      GAATCCGTGT CGTGTTACGG GTGGTGGTGG TCACACGGCG TGTTCCGGCA

1701  GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT
      CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA

                      BfrI
                      ~~~~~~~
                      AflII
                      ~~~~~~~
1751  GGACGCAGAT GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT
      CCTGCGTCTA CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA

                                                      HpaI
                                                      ~
1801  GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT TGCGGTGCTG
      CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA ACGCCACGAC

      HpaI
      ~~~~~
1851  TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG
      AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC

1901  CGCCACCAGA CATAATAGCT GACAGACTAA CAGACTGTTC CTTTCCATGG
      GCGGTGGTCT GTATTATCGA CTGTCTGATT GTCTGACAAG GAAAGGTACC

  +3                          SEQ ID NO: 5 ——→ M  D  A
                          SalI
                          ~~~~~~
1951  GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA
      CAGAAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCGTTA GTACCTACGT

  +3  M  K  R  G   L  C  C   V  L  L   L  C  G  A   V  F  V
2001  ATGAAGAGAG GGCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT
      TACTTCTCTC CCGAGACGAC ACACGACGAC GACACACCTC GTCAGAAGCA
```

FIG. 3D

```
   +3   S  P  S  A  S  Y  Q  V  R  N  S  T  G  L  Y  H
                NheI                                PmlI
                ~~~~~~                              ~~~~
           Eco47III                                 PmaCI
            ~~~~~~~                                 ~~~~
             Afe  I      SexAI                      BbrPI
            ~~~~~~~     ~~~~~~~~                    ~~~~
2051  TTCGCCCAGC GCTAGCTACC AGGTGCGCAA CAGCACCGGC CTGTACCACG
      AAGCGGGTCG CGATCGATGG TCCACGCGTT GTCGTGGCCG GACATGGTGC

   +3 V  T  N  D  C  P  N  S  S  I  V  Y  E  A  A  D  A
      PmlI
      ~~
      PmaCI
      ~~
      BbrPI
      ~~
2101  TGACCAACGA CTGCCCCAAC AGCAGCATCG TGTACGAGGC CGCCGACGCC
      ACTGGTTGCT GACGGGGTTG TCGTCGTAGC ACATGCTCCG GCGGCTGCGG

   +3  I  L  H  T  P  G  C  V  P  C  V  R  E  G  N  A  S
2151  ATCCTGCACA CCCCCGGCTG CGTGCCCTGC GTGCGCGAGG GCAACGCCAG
      TAGGACGTGT GGGGGCCGAC GCACGGGACG CACGCGCTCC CGTTGCGGTC

   +3   R  C  W   V  A  M  T   P  T  V   A  T  R   D  G  K
2201  CCGCTGCTGG GTGGCCATGA CCCCCACCGT GGCCACCCGC GACGGCAAGC
      GGCGACGACC CACCGGTACT GGGGGTGGCA CCGGTGGGCG CTGCCGTTCG

   +3 L  P  A  T  Q  L  R  R  H  I  D  L  L  V  G  S  A
                                                  DraIII
                                                  ~
2251  TGCCCGCCAC CCAGCTGCGC CGCCACATCG ACCTGCTGGT GGGCAGCGCC
      ACGGGCGGTG GGTCGACGCG GCGGTGTAGC TGGACGACCA CCCGTCGCGG

   +3  T  L  C  S  A  L  Y  V  G  D  L  C  G  S  V  F  L
       DraIII
       ~~~~~~~~
2301  ACCCTGTGCA GCGCCCTGTA CGTGGGCGAC CTGTGCGGCA GCGTGTTCCT
      TGGGACACGT CGCGGGACAT GCACCCGCTG GACACGCCGT CGCACAAGGA

   +3   V  G  Q   L  F  T  F   S  P  R   R  H  W   T  T  Q
2351  GGTGGGCCAG CTGTTCACCT TCAGCCCCCG CCGCCACTGG ACCACCCAGG
      CCACCCGGTC GACAAGTGGA AGTCGGGGGC GGCGGTGACC TGGTGGGTCC

   +3 G  C  N  C   S  I  Y   P  G  H  I   T  G  H   R  M  A
2401  GCTGCAACTG CAGCATCTAC CCCGGCCACA TCACCGGCCA CCGCATGGCC
      CGACGTTGAC GTCGTAGATG GGGCCGGTGT AGTGGCCGGT GGCGTACCGG

   +3  W  D  M  M   M  N  W   S  P  T   T  M  E  N   I  T  S
2451  TGGGACATGA TGATGAACTG GAGCCCCACC ACCATGGAGA ACATCACATC
      ACCCTGTACT ACTACTTGAC CTCGGGGTGG TGGTACCTCT TGTAGTGTAG

   +3   G  F  L   G  P  L  L   V  L  Q   A  G  F   F  L  L
                  PpuMI
                  ~~~~~~~~
2501  AGGATTCCTA GGACCCCTGC TCGTGTTACA GGCGGGGTTT TTCTTGTTGA
      TCCTAAGGAT CCTGGGGACG AGCACAATGT CCGCCCCAAA AAGAACAACT
```

FIG. 3E

```
   +3 T  R  I  L  T  P  Q  S  L  D  S  W  W  T  S  L
2551  CAAGAATCCT CACAATACCG CAGAGTCTAG ACTCGTGGTG GACTTCTCTC
      GTTCTTAGGA GTGTTATGGC GTCTCAGATC TGAGCACCAC CTGAAGAGAG

   +3 N  F  L  G  G  S  P  V  C  L  G  Q  N  S  Q  S  P
2601  AATTTTCTAG GGGGATCTCC CGTGTGTCTT GGCCAAAATT CGCAGTCCCC
      TTAAAAGATC CCCCTAGAGG GCACACAGAA CCGGTTTTAA GCGTCAGGGG

   +3  T  S  N  H  S  P  T  S  C  P  P  I  C  P  G  Y
2651  AACCTCCAAT CACTCACCAA CCTCCTGTCC TCCAATTTGT CCTGGTTATC
      TTGGAGGTTA GTGAGTGGTT GGAGGACAGG AGGTTAAACA GGACCAATAG

   +3 R  W  M  C  L  R  R  F  I  I  F  L  F  I  L  L  L
2701  GCTGGATGTG TCTGCGGCGT TTTATCATAT TCCTCTTCAT CCTGCTGCTA
      CGACCTACAC AGACGCCGCA AAATAGTATA AGGAGAAGTA GGACGACGAT

   +3 C  L  I  F  L  L  V  L  L  D  Y  Q  G  M  L  P  V
2751  TGCCTCATCT TCTTATTGGT TCTTCTGGAT TATCAAGGTA TGTTGCCCGT
      ACGGAGTAGA AGAATAACCA AGAAGACCTA ATAGTTCCAT ACAACGGGCA

   +3  C  P  L  I  P  G  S  T  T  T  S  T  G  P  C  K
                                                 BstAP I
                                                 ~~~~~~
2801  TTGTCCTCTA ATTCCAGGAT CAACAACAAC CAGTACGGGA CCATGCAAAA
      AACAGGAGAT TAAGGTCCTA GTTGTTGTTG GTCATGCCCT GGTACGTTTT

   +3 T  C  T  T  P  A  Q  G  N  S  M  F  P  S  C  C  C
      BstAP I        EcoNI
      ~~~~~          ~~~~~~~~~~~~
2851  CCTGCACGAC TCCTGCTCAA GGCAACTCTA TGTTTCCCTC ATGTTGCTGT
      GGACGTGCTG AGGACGAGTT CCGTTGAGAT ACAAAGGGAG TACAACGACA

   +3 T  K  P  T  D  G  N  C  T  C  I  P  I  P  S  S  W
2901  ACAAAACCTA CGGATGGAAA TTGCACCTGT ATTCCCATCC CATCGTCCTG
      TGTTTTGGAT GCCTACCTTT AACGTGGACA TAAGGGTAGG GTAGCAGGAC

   +3  A  F  A  K  Y  L  W  E  W  A  S  V  R  F  S  W
2951  GGCTTTCGCA AAATACCTAT GGGAGTGGGC CTCAGTCCGT TTCTCTTGGC
      CCGAAAGCGT TTTATGGATA CCCTCACCCG GAGTCAGGCA AAGAGAACCG

   +3 L  S  L  L  V  P  F  V  Q  W  F  V  G  L  S  P  T
3001  TCAGTTTACT AGTGCCATTT GTTCAGTGGT TCGTAGGGCT TTCCCCCACT
      AGTCAAATGA TCACGGTAAA CAAGTCACCA AGCATCCCGA AAGGGGGTGA

   +3 V  W  L  S  A  I  W  M  M  W  Y  W  G  P  S  L  Y
3051  GTTTGGCTTT CAGCTATATG GATGATGTGG TATTGGGGGC CAAGTCTGTA
      CAAACCGAAA GTCGATATAC CTACTACACC ATAACCCCCG GTTCAGACAT

   +3  S  I  V  S  P  F  I  P  L  L  P  I  F  F  C  L
3101  CAGCATCGTG AGTCCCTTTA TACCGCTGTT ACCAATTTTC TTTTGTCTCT
      GTCGTAGCAC TCAGGGAAAT ATGGCGACAA TGGTTAAAAG AAAACAGAGA

   +3 W  V  Y  I  *
       BstZ17 I              XhoI
       ~~~~~~                ~~~~~~
       Bst1107I              PaeR7I              AscI
       ~~~~~~                ~~~~~~              ~~~~~~~~
3151  GGGTATACAT TTAAGAATTC AGACTCGAGC AAGTCTAGAA AGGCGCGCCA
      CCCATATGTA AATTCTTAAG TCTGAGCTCG TTCAGATCTT TCCGCGCGGT
```

FIG. 3F

```
           EcoRV     BamHI    MluI             BclI
           ~~~~~~  ~~~~~~~  ~~~~~~~          ~~~~~~
3201  AGATATCAAG GATCCACTAC GCGTTAGAGC TCGCTGATCA GCCTCGACTG
      TCTATAGTTC CTAGGTGATG CGCAATCTCG AGCGACTAGT CGGAGCTGAC

3251  TGCCTTCTAG TTGCCAGCCA TCTGTTGTTT GCCCCTCCCC CGTGCCTTCC
      ACGGAAGATC AACGGTCGGT AGACAACAAA CGGGGAGGGG GCACGGAAGG

3301  TTGACCCTGG AAGGTGCCAC TCCCACTGTC CTTTCCTAAT AAAATGAGGA
      AACTGGGACC TTCCACGGTG AGGGTGACAG GAAAGGATTA TTTTACTCCT

3351  AATTGCATCG CATTGTCTGA GTAGGTGTCA TTCTATTCTG GGGGGTGGGG
      TTAACGTAGC GTAACAGACT CATCCACAGT AAGATAAGAC CCCCCACCCC

3401  TGGGGCAGGA CAGCAAGGGG GAGGATTGGG AAGACAATAG CAGGCATGCT
      ACCCCGTCCT GTCGTTCCCC CTCCTAACCC TTCTGTTATC GTCCGTACGA

3451  GGGGAGCTCT TCCGCTTCCT CGCTCACTGA CTCGCTGCGC TCGGTCGTTC
      CCCCTCGAGA AGGCGAAGGA GCGAGTGACT GAGCGACGCG AGCCAGCAAG

3501  GGCTGCGGCG AGCGGTATCA GCTCACTCAA AGGCGGTAAT ACGGTTATCC
      CCGACGCCGC TCGCCATAGT CGAGTGAGTT TCCGCCATTA TGCCAATAGG

                                   Pci  I
                                  ~~~~~~~
3551  ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA
      TGTCTTAGTC CCCTATTGCG TCCTTTCTTG TACACTCGTT TTCCGGTCGT

3601  AAAGGCCAGG AACCGTAAAA AGGCCGCGTT GCTGGCGTTT TTCCATAGGC
      TTTCCGGTCC TTGGCATTTT TCCGGCGCAA CGACCGCAAA AAGGTATCCG

3651  TCCGCCCCCC TGACGAGCAT CACAAAAATC GACGCTCAAG TCAGAGGTGG
      AGGCGGGGGG ACTGCTCGTA GTGTTTTTAG CTGCGAGTTC AGTCTCCACC

3701  CGAAACCCGA CAGGACTATA AAGATACCAG GCGTTTCCCC CTGGAAGCTC
      GCTTTGGGCT GTCCTGATAT TTCTATGGTC CGCAAAGGGG GACCTTCGAG

3751  CCTCGTGCGC TCTCCTGTTC CGACCCTGCC GCTTACCGGA TACCTGTCCG
      GGAGCACGCG AGAGGACAAG GCTGGGACGG CGAATGGCCT ATGGACAGGC

3801  CCTTTCTCCC TTCGGGAAGC GTGGCGCTTT CTCAATGCTC ACGCTGTAGG
      GGAAAGAGGG AAGCCCTTCG CACCGCGAAA GAGTTACGAG TGCGACATCC

3851  TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA
      ATAGAGTCAA GCCACATCCA GCAAGCGAGG TTCGACCCGA CACACGTGCT

3901  ACCCCCCGTT CAGCCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG
      TGGGGGGCAA GTCGGGCTGG CGACGCGGAA TAGGCCATTG ATAGCAGAAC

3951  AGTCCAACCC GGTAAGACAC GACTTATCGC CACTGGCAGC AGCCACTGGT
      TCAGGTTGGG CCATTCTGTG CTGAATAGCG GTGACCGTCG TCGGTGACCA

4001  AACAGGATTA GCAGAGCGAG GTATGTAGGC GGTGCTACAG AGTTCTTGAA
      TTGTCCTAAT CGTCTCGCTC CATACATCCG CCACGATGTC TCAAGAACTT

4051  GTGGTGGCCT AACTACGGCT ACACTAGAAG GACAGTATTT GGTATCTGCG
      CACCACCGGA TTGATGCCGA TGTGATCTTC CTGTCATAAA CCATAGACGC
```

FIG. 3G

```
4101  CTCTGCTGAA GCC  ACC TTCGGAAAAA GAGTTGGTAG CTCTTGATCC
      GAGACGACTT CGGTCAATGG AAGCCTTTTT CTCAACCATC GAGAACTAGG

4151  GGCAAACAAA CCACCGCTGG TAGCGGTGGT TTTTTTGTTT GCAAGCAGCA
      CCGTTTGTTT GGTGGCGACC ATCGCCACCA AAAAAACAAA CGTTCGTCGT

4201  GATTACGCGC AGAAAAAAAG GATCTCAAGA AGATCCTTTG ATCTTTTCTA
      CTAATGCGCG TCTTTTTTTC CTAGAGTTCT TCTAGGAAAC TAGAAAAGAT

4251  CGGGGTCTGA CGCTCAGTGG AACGAAAACT CACGTTAAGG GATTTTGGTC
      GCCCCAGACT GCGAGTCACC TTGCTTTTGA GTGCAATTCC CTAAAACCAG

4301  ATGAGATTAT CAAAAAGGAT CTTCACCTAG ATCCTTTTAA ATTAAAAATG
      TACTCTAATA GTTTTTCCTA GAAGTGGATC TAGGAAAATT TAATTTTTAC

4351  AAGTTTTAAA TCAATCTAAA GTATATATGA GTAAACTTGG TCTGACAGTT
      TTCAAAATTT AGTTAGATTT CATATATACT CATTTGAACC AGACTGTCAA

4401  ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTTCGT
      TGGTTACGAA TTAGTCACTC CGTGGATAGA GTCGCTAGAC AGATAAAGCA

                      Eam1105I
                      ~~~~~~~~~~~~
                        AspEI
                      ~~~~~~~~~~~~
4451  TCATCCATAG TTGCCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA
      AGTAGGTATC AACGGACTGA GGGGCAGCAC ATCTATTGAT GCTATGCCCT

4501  GGGCTTACCA TCTGGCCCCA GTGCTGCAAT GATACCGCGA GACCCACGCT
      CCCGAATGGT AGACCGGGGT CACGACGTTA CTATGGCGCT CTGGGTGCGA
                                              BsaI
                                            ~~~~~~~
4551  CACCGGCTCC AGATTTATCA GCAATAAACC AGCCAGCCGG AAGGGCCGAG
      GTGGCCGAGG TCTAAATAGT CGTTATTTGG TCGGTCGGCC TTCCCGGCTC

4601  CGCAGAAGTG GTCCTGCAAC TTTATCCGCC TCCATCCAGT CTATTAATTG
      GCGTCTTCAC CAGGACGTTG AAATAGGCGG AGGTAGGTCA GATAATTAAC

4651  TTGCCGGGAA GCTAGAGTAA GTAGTTCGCC AGTTAATAGT TTGCGCAACG
      AACGGCCCTT CGATCTCATT CATCAAGCGG TCAATTATCA AACGCGTTGC

4701  TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC GTTTGGTATG
      AACAACGGTA ACGATGTCCG TAGCACCACA GTGCGAGCAG CAAACCATAC

4751  GCTTCATTCA GCTCCGGTTC CCAACGATCA AGGCGAGTTA CATGATCCCC
      CGAAGTAAGT CGAGGCCAAG GGTTGCTAGT TCCGCTCAAT GTACTAGGGG

                                              PvuI
                                            ~~~~~~~
4801  CATGTTGTGC AAAAAAGCGG TTAGCTCCTT CGGTCCTCCG ATCGTTGTCA
      GTACAACACG TTTTTTCGCC AATCGAGGAA GCCAGGAGGC TAGCAACAGT

4851  GAAGTAAGTT GGCCGCAGTG TTATCACTCA TGGTTATGGC AGCACTGCAT
      CTTCATTCAA CCGGCGTCAC AATAGTGAGT ACCAATACCG TCGTGACGTA

4901  AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTTCTG TGACTGGTGA
      TTAAGAGAAT GACAGTACGG TAGGCATTCT ACGAAAAGAC ACTGACCACT
```

FIG. 3H

```
                                                          BcgI
                                                      ~~~~~~~~~~~~
4951  GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGGCGA CCGAGTTGCT
      CATGAGTTGG TTCAGTAAGA CTCTTATCAC ATACGCCGCT GGCTCAACGA

5001  CTTGCCCGGC GTCAATACGG GATAATACCG CGCCACATAG CAGAACTTTA
      GAACGGGCCG CAGTTATGCC CTATTATGGC GCGGTGTATC GTCTTGAAAT

                        XmnI
                    ~~~~~~~~~~~~
                       Asp700
                    ~~~~~~~~~~~~
5051  AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC TCTCAAGGAT
      TTTCACGAGT AGTAACCTTT TGCAAGAAGC CCCGCTTTTG AGAGTTCCTA

5101  CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCCAACT
      GAATGGCGAC AACTCTAGGT CAAGCTACAT TGGGTGAGCA CGTGGGTTGA

5151  GATCTTCAGC ATCTTTTACT TTCACCAGCG TTTCTGGGTG AGCAAAAACA
      CTAGAAGTCG TAGAAAATGA AAGTGGTCGC AAAGACCCAC TCGTTTTTGT

5201  GGAAGGCAAA ATGCCGCAAA AAAGGGAATA AGGGCGACAC GGAAATGTTG
      CCTTCCGTTT TACGGCGTTT TTTCCCTTAT TCCCGCTGTG CCTTTACAAC

                                  SspI
                               ~~~~~~
5251  AATACTCATA CTCTTCCTTT TTCAATATTA TTGAAGCATT TATCAGGGTT
      TTATGAGTAT GAGAAGGAAA AAGTTATAAT AACTTCGTAA ATAGTCCCAA

5301  ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA
      TAACAGAGTA CTCGCCTATG TATAAACTTA CATAAATCTT TTTATTTGTT

5351  ATAGGGGTTC CGCGCACATT TCCCCGAAAA GTGCCACCTG ACGTCTAAGA
      TATCCCCAAG GCGCGTGTAA AGGGGCTTTT CACGGTGGAC TGCAGATTCT

5401  AACCATTATT ATCATGACAT TAACCTATAA AAATAGGCGT ATCACGAGGC
      TTGGTAATAA TAGTACTGTA ATTGGATATT TTTATCCGCA TAGTGCTCCG

5451  CCTTTCGTC
      GGAAAGCAG
```

FIG. 3I

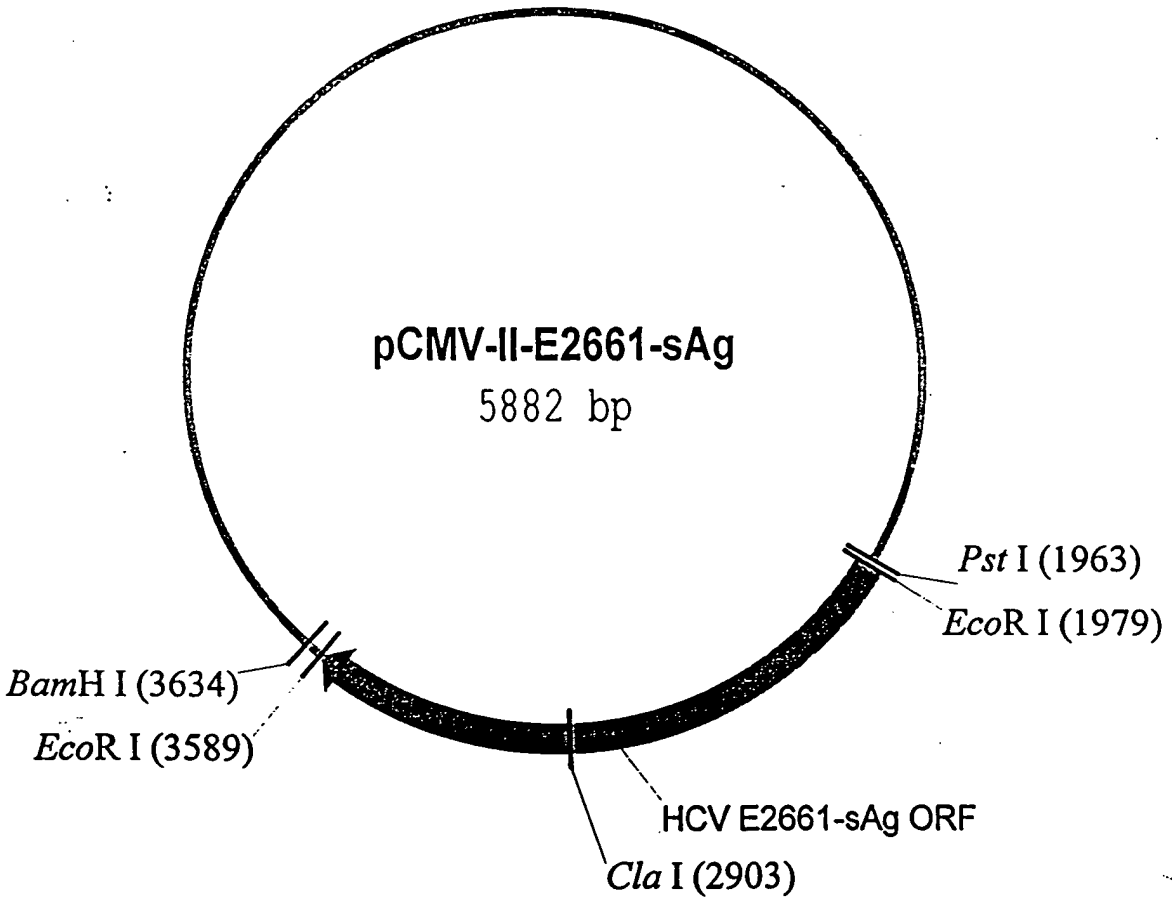
pCMV-II-E2661-sAg
5882 bp
Pst I (1963)
*Eco*R I (1979)
*Bam*H I (3634)
*Eco*R I (3589)
HCV E2661-sAg ORF
Cla I (2903)

FIG. 4A

pCMV-II-E2661-sAg

SEQ ID NO: 6

```
1    TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT
     AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC CTCTGCCAGT GTCGAACAGA CATTCGCCTA

81   GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG CGGCATCAGA
     CGGCCCTCGT CTGTTCGGGC AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC GCCGTAGTCT

161  GCAGATTGTA CTGAGAGTGC ACCATATGAA GCTTTTTGCA AAAGCCTAGG CCTCCAAAAA AGCCTCCTCA CTACTTCTGG
     CGTCTAACAT GACTCTCACG TGGTATACTT CGAAAAACGT TTTCGGATCC GGAGGTTTTT TCGGAGGAGT GATGAAGACC

241  AATAGCTCAG AGGCCGAGGC GGCCTCGGCC TCTGCATAAA TAAAAAAAAT TAGTCAGCCA TGGGGCGGAG AATGGGCGGA
     TTATCGAGTC TCCGGCTCCG CCGGAGCCGG AGACGTATTT ATTTTTTTTA ATCAGTCGGT ACCCCGCCTC TTACCCGCCT

321  ACTGGGCGGG GAGGGAATTA TTGGCTATTG GCCATTGCAT ACGTTGTATC TATATCATAA TATGTACATT TATATTGGCT
     TGACCCGCCC CTCCCTTAAT AACCGATAAC CGGTAACGTA TGCAACATAG ATATAGTATT ATACATGTAA ATATAACCGA

401  CATGTCCAAT ATGACCGCCA TGTTGACATT GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT
     GTACAGGTTA TACTGGCGGT ACAACTGTAA CTAATAACTG ATCAATAATT ATCATTAGTT AATGCCCCAG TAATCAAGTA

481  AGCCCATATA TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC CCCGCCCATT
     TCGGGTATAT ACCTCAAGGC GCAATGTATT GAATGCCATT TACCGGGCGG ACCGACTGGC GGGTTGCTGG GGGCGGGTAA

561  GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC ATTGACGTCA ATGGGTGGAG TATTTACGGT
     CTGCAGTTAT TACTGCATAC AAGGGTATCA TTGCGGTTAT CCCTGAAAGG TAACTGCAGT TACCCACCTC ATAAATGCCA

641  AAACTGCCCA CTTGGCAGTA CATCAAGTGT ATCATATGCC AAGTCCGCCC CCTATTGACG TCAATGACGG TAAATGGCCC
     TTTGACGGGT GAACCGTCAT GTAGTTCACA TAGTATACGG TTCAGGCGGG GGATAACTGC AGTTACTGCC ATTTACCGGG

721  GCCTGGCATT ATGCCCAGTA CATGACCTTA CGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA TCGCTATTAC
     CGGACCGTAA TACGGGTCAT GTACTGGAAT GCCCTGAAAG GATGAACCGT CATGTAGATG CATAATCAGT AGCGATAATG

801  CATGGTGATG CGGTTTTGGC AGTACACCAA TGGGCGTGGA TAGCGGTTTG ACTCACGGGG ATTTCCAAGT CTCCACCCCA
     GTACCACTAC GCCAAAACCG TCATGTGGTT ACCCGCACCT ATCGCCAAAC TGAGTGCCCC TAAAGGTTCA GAGGTGGGGT

881  TTGACGTCAA TGGGAGTTTG TTTTGGCACC AAAATCAACG GGACTTTCCA AAATGTCGTA ATAACCCCGC CCCGTTGACG
     AACTGCAGTT ACCCTCAAAC AAAACCGTGG TTTTAGTTGC CCTGAAAGGT TTTACAGCAT TATTGGGGCG GGGCAACTGC

961  CAAATGGGCG GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCG TTTAGTGAAC CGTCAGATCG CCTGGAGACG
     GTTTACCCGC CATCCGCACA TGCCACCCTC CAGATATATT CGTCTCGAGC AAATCACTTG GCAGTCTAGC GGACCTCTGC

1041 CCATCCACGC TGTTTTGACC TCCATAGAAG ACACCGGGAC CGATCCAGCC TCCGCGGCCG GGAACGGTGC ATTGGAACGC
     GGTAGGTGCG ACAAAACTGG AGGTATCTTC TGTGGCCCTG GCTAGGTCGG AGGCGCCGGC CCTTGCCACG TAACCTTGCG

1121 GGATTCCCCG TGCCAAGAGT GACGTAAGTA CCGCCTATAG ACTCTATAGG CACACCCCTT TGGCTCTTAT GCATGCTATA
     CCTAAGGGGC ACGGTTCTCA CTGCATTCAT GGCGGATATC TGAGATATCC GTGTGGGGAA ACCGAGAATA CGTACGATAT

1201 CTGTTTTTGG CTTGGGGCCT ATACACCCCC GCTCCTTATG CTATAGGTGA TGGTATAGCT TAGCCTATAG GTGTGGGTTA
     GACAAAAACC GAACCCCGGA TATGTGGGGG CGAGGAATAC GATATCCACT ACCATATCGA ATCGGATATC CACACCCAAT

1281 TTGACCATTA TTGACCACTC CCCTATTGGT GACGATACTT TCCATTACTA ATCCATAACA TGGCTCTTTG CCACAACTAT
     AACTGGTAAT AACTGGTGAG GGGATAACCA CTGCTATGAA AGGTAATGAT TAGGTATTGT ACCGAGAAAC GGTGTTGATA

1361 CTCTATTGGC TATATGCCAA TACTCTGTCC TTCAGAGACT GACACGGACT CTGTATTTTT ACAGGATGGG GTCCATTTAT
     GAGATAACCG ATATACGGTT ATGAGACAGG AAGTCTCTGA CTGTGCCTGA GACATAAAAA TGTCCTACCC CAGGTAAATA

1441 TATTTACAAA TTCACATATA CAACAACGCC GTCCCCCGTG CCCGCAGTTT TTATTAAACA TAGCGTGGGA TCTCCGACAT
     ATAAATGTTT AAGTGTATAT GTTGTTGCGG CAGGGGGCAC GGGCGTCAAA AATAATTTGT ATCGCACCCT AGAGGCTGTA
```

FIG. 4B

```
1521  CTCGGGTACG TGTTCCGGAC ATGGGCTCTT CTCCGGTAGC GGCGGAGCTT CCACATCCGA GCCCTGGTCC CATCCGTCCA
      GAGCCCATGC ACAAGGCCTG TACCCGAGAA GAGGCCATCG CCGCCTCGAA GGTGTAGGCT CGGGACCAGG GTAGGCAGGT

1601  GCGGCTCATG GTCGCTCGGC AGCTCCTTGC TCCTAACAGT GGAGGCCAGA CTTAGGCACA GCACAATGCC CACCACCACC
      CGCCGAGTAC CAGCGAGCCG TCGAGGAACG AGGATTGTCA CCTCCGGTCT GAATCCGTGT CGTGTTACGG GTGGTGGTGG

1681  AGTGTGCCGC ACAAGGCCGT GGCGGTAGGG TATGTGTCTG AAAATGAGCT CGGAGATTGG GCTCGCACCT GGACGCAGAT
      TCACACGGCG TGTTCCGGCA CCGCCATCCC ATACACAGAC TTTTACTCGA GCCTCTAACC CGAGCGTGGA CCTGCGTCTA

1761  GGAAGACTTA AGGCAGCGGC AGAAGAAGAT GCAGGCAGCT GAGTTGTTGT ATTCTGATAA GAGTCAGAGG TAACTCCCGT
      CCTTCTGAAT TCCGTCGCCG TCTTCTTCTA CGTCCGTCGA CTCAACAACA TAAGACTATT CTCAGTCTCC ATTGAGGGCA

1841  TGCGGTGCTG TTAACGGTGG AGGGCAGTGT AGTCTGAGCA GTACTCGTTG CTGCCGCGCG CGCCACCAGA CATAATAGCT
      ACGCCACGAC AATTGCCACC TCCCGTCACA TCAGACTCGT CATGAGCAAC GACGGCGCGC GCGGTGGTCT GTATTATCGA

                                                       SEQ ID NO: 7 ------------> M  D  A
  +3                                        PstI              EcoRI
                                          ~~~~~~~           ~~~~~~~
1921  GACAGACTAA CAGACTGTTC CTTTCCATGG GTCTTTTCTG CAGTCACCGT CGTCGACGAA TTCAAGCAAT CATGGATGCA
      CTGTCTGATT GTCTGACAAG GAAAGGTACC CAGAAAAGAC GTCAGTGGCA GCAGCTGCTT AAGTTCGTTA GTACCTACGT

  +3  M  K  R  G   L  C  C   V  L  L   L  C  G  A   V  F  V   S  P  S   A  S  E  T   H  V  T
2001  ATGAAGAGAG GCTCTGCTG TGTGCTGCTG CTGTGTGGAG CAGTCTTCGT TTCGCCCAGC GCTAGCGAAA CCCACGTCAC
      TACTTCTCTC CCGAGACGAC ACACGACGAC GACACACCTC GTCAGAAGCA AAGCGGGTCG CGATCGCTTT GGGTGCAGTG

  +3   G  G  S   A  G  H  T   V  S  G   F  V  S   L  L  A  P   G  A  K   Q  N  V   Q  L  I
2081  CGGGGGAAGT GCCGGCCACA CTGTGTCTGG ATTTGTTAGC CTCCTCGCAC CAGGCGCCAA GCAGAACGTC CAGCTGATCA
      GCCCCCTTCA CGGCCGGTGT GACACAGACC TAAACAATCG GAGGAGCGTG GTCCGCGGTT CGTCTTGCAG GTCGACTAGT

  +3  N  T  N  G   S  W  H   L  N  S  T   A  L  N   C  N  D   S  L  N  T   G  W  L   A  G  L
2161  ACACCAACGG CAGTTGGCAC CTCAATAGCA CGGCCCTGAA CTGCAATGAT AGCCTCAACA CCGGCTGGTT GGCAGGGCTT
      TGTGGTTGCC GTCAACCGTG GAGTTATCGT GCCGGGACTT GACGTTACTA TCGGAGTTGT GGCCGACCAA CCGTCCCGAA

  +3   F  Y  H  H   K  F  N   S  S  G   C  P  E  R   L  A  S   C  R  P   L  T  D  F   D  Q  G
2241  TTCTATCACC ACAAGTTCAA CTCTTCAGGC TGTCCTGAGA GGCTAGCCAG CTGCCGACCC CTTACCGATT TTGACCAGGG
      AAGATAGTGG TGTTCAAGTT GAGAAGTCCG ACAGGACTCT CCGATCGGTC GACGGCTGGG GAATGGCTAA AACTGGTCCC

  +3   W  G  P   I  S  Y  A   N  G  S   G  P  D   Q  R  P  Y   C  W  H   Y  P  P   K  P  C
2321  CTGGGGCCCT ATCAGTTATG CCAACGGAAG CGGCCCCGAC CAGCGCCCCT ACTGCTGGCA CTACCCCCCA AAACCTTGCG
      GACCCCGGGA TAGTCAATAC GGTTGCCTTC GCCGGGGCTG GTCGCGGGGA TGACGACCGT GATGGGGGGT TTTGGAACGC

  +3  G  I  V  P   A  K  S   V  C  G  P   V  Y  C   F  T  P   S  P  V  V   V  G  T   T  D  R
2401  GTATTGTGCC CGCGAAGAGT GTGTGTGGTC CGGTATATTG CTTCACTCCC AGCCCCGTGG TGGTGGGAAC GACCGACAGG
      CATAACACGG GCGCTTCTCA CACACACCAG GCCATATAAC GAAGTGAGGG TCGGGGCACC ACCACCCTTG CTGGCTGTCC

  +3   S  G  A  P   T  Y  S   W  G  E   N  D  T  D   V  F  V   L  N  N   T  R  P  P   L  G  N
2481  TCGGGCGCGC CCACCTACAG CTGGGGTGAA AATGATACGG ACGTCTTCGT CCTTAACAAT ACCAGGCCAC CGCTGGGCAA
      AGCCCGCGCG GGTGGATGTC GACCCCACTT TTACTATGCC TGCAGAAGCA GGAATTGTTA TGGTCCGGTG GCGACCCGTT

  +3   W  F  G   C  T  W  M   N  S  T   G  F  T   K  V  C  G   A  P  P   C  V  I   G  G  A
2561  TTGGTTCGGT TGTACCTGGA TGAACTCAAC TGGATTCACC AAAGTGTGCG GAGCGCCTCC TTGTGTCATC GGAGGGGCGG
      AACCAAGCCA ACATGGACCT ACTTGAGTTG ACCTAAGTGG TTTCACACGC CTCGCGGAGG AACACAGTAG CCTCCCCGCC

  +3  G  N  N  T   L  H  C   P  T  D  C   F  R  K   H  P  D   A  T  Y  S   R  C  G   S  G  P
2641  GCAACAACAC CCTGCACTGC CCCACTGATT GCTTCCGCAA GCATCCGGAC GCCACATACT CTCGGTGCGG CTCCGGTCCC
      CGTTGTTGTG GGACGTGACG GGGTGACTAA CGAAGGCGTT CGTAGGCCTG CGGTGTATGA GAGCCACGCC GAGGCCAGGG
```

FIG. 4C

```
   +3  W  I  T  P  R  C  L  V  D  Y  P  Y  R  L  W  H  Y  P  C  T  I  N  Y  T  I  F  K
2721  TGGATCACAC CCAGGTGCCT GGTCGACTAC CCGTATAGGC TTTGGCATTA TCCTTGTACC ATCAACTACA CCATATTTAA
      ACCTAGTGTG GGTCCACGGA CCAGCTGATG GGCATATCCG AAACCGTAAT AGGAACATGG TAGTTGATGT GGTATAAATT

   +3  I  R  M  Y  V  G  G  V  E  H  R  L  E  A  A  C  N  W  T  R  G  E  R  C  D  L
2801  AATCAGGATG TACGTGGGAG GGGTCGAACA CAGGCTGGAA GCTGCCTGCA ACTGGACGCG GGGCGAACGT TGCGATCTGG
      TTAGTCCTAC ATGCACCCTC CCCAGCTTGT GTCCGACCTT CGACGGACGT TGACCTGCGC CCCGCTTGCA ACGCTAGACC

   +3  E  D  R  D  R  S  E  I  D  M  E  N  I  T  S  G  F  L  G  P  L  L  V  L  Q  A  G
                             ClaI
                             ~~~~~~
2881  AAGATAGGGA CAGGTCCGAG ATCGATATGG AGAACATCAC ATCAGGATTC CTAGGACCCC TGCTCGTGTT ACAGGCGGGG
      TTCTATCCCT GTCCAGGCTC TAGCTATACC TCTTGTAGTG TAGTCCTAAG GATCCTGGGG ACGAGCACAA TGTCCGCCCC

   +3  F  F  L  L  T  R  I  L  T  I  P  Q  S  L  D  S  W  W  T  S  L  N  F  L  G  G  S
2961  TTTTTCTTGT TGACAAGAAT CCTCACAATA CCGCAGAGTC TAGACTCGTG GTGGACTTCT CTCAATTTTC TAGGGGGATC
      AAAAAGAACA ACTGTTCTTA GGAGTGTTAT GGCGTCTCAG ATCTGAGCAC CACCTGAAGA GAGTTAAAAG ATCCCCCTAG

   +3  P  V  C  L  G  Q  N  S  Q  S  P  T  S  N  H  S  P  T  S  C  P  P  I  C  P  G
3041  TCCCGTGTGT CTTGGCCAAA ATTCGCAGTC CCCAACCTCC AATCACTCAC CAACCTCCTG TCCTCCAATT TGTCCTGGTT
      AGGGCACACA GAACCGGTTT TAAGCGTCAG GGGTTGGAGG TTAGTGAGTG GTTGGAGGAC AGGAGGTTAA ACAGGACCAA

   +3  Y  R  W  M  C  L  R  R  F  I  I  F  L  F  I  L  L  L  C  L  I  F  L  L  V  L  L
3121  ATCGCTGGAT GTGTCTGCGG CGTTTTATCA TATTCCTCTT CATCCTGCTG CTATGCCTCA TCTTCTTATT GGTTCTTCTG
      TAGCGACCTA CACAGACGCC GCAAAATAGT ATAAGGAGAA GTAGGACGAC GATACGGAGT AGAAGAATAA CCAAGAAGAC

   +3  D  Y  Q  G  M  L  P  V  C  P  L  I  P  G  S  T  T  T  S  T  G  P  C  K  T  C  T
3201  GATTATCAAG GTATGTTGCC CGTTTGTCCT CTAATTCCAG GATCAACAAC AACCAGTACG GGACCATGCA AAACCTGCAC
      CTAATAGTTC CATACAACGG GCAAACAGGA GATTAAGGTC CTAGTTGTTG TTGGTCATGC CCTGGTACGT TTTGGACGTG

   +3  T  P  A  Q  G  N  S  M  F  P  S  C  C  C  T  K  P  T  D  G  N  C  T  C  I  P
3281  GACTCCTGCT CAAGGCAACT CTATGTTTCC CTCATGTTGC TGTACAAAAC CTACGGATGG AAATTGCACC TGTATTCCCA
      CTGAGGACGA GTTCCGTTGA GATACAAAGG GAGTACAACG ACATGTTTTG GATGCCTACC TTTAACGTGG ACATAAGGGT

   +3  I  P  S  S  W  A  F  A  K  Y  L  W  E  W  A  S  V  R  F  S  W  L  S  L  L  V  P
3361  TCCCATCGTC CTGGGCTTTC GCAAAATACC TATGGGAGTG GGCCTCAGTC CGTTTCTCTT GGCTCAGTTT ACTAGTGCCA
      AGGGTAGCAG GACCCGAAAG CGTTTTATGG ATACCCTCAC CCGGAGTCAG GCAAAGAGAA CCGAGTCAAA TGATCACGGT

   +3  F  V  Q  W  F  V  G  L  S  P  T  V  W  L  S  A  I  W  M  M  W  Y  W  G  P  S  L
3441  TTTGTTCAGT GGTTCGTAGG GCTTTCCCCC ACTGTTTGGC TTTCAGCTAT ATGGATGATG TGGTATTGGG GGCCAAGTCT
      AAACAAGTCA CCAAGCATCC CGAAAGGGGG TGACAAACCG AAAGTCGATA TACCTACTAC ACCATAACCC CCGGTTCAGA

   +3  Y  S  I  V  S  P  F  I  P  L  L  P  I  F  F  C  L  W  V  Y  I  *
                                                                  EcoRI
                                                                  ~~~~~~~~
3521  GTACAGCATC GTGAGTCCCT TTATACCGCT GTTACCAATT TTCTTTTGTC TCTGGGTATA CATTTAAGAA TTCAGACTCG
      CATGTCGTAG CACTCAGGGA AATATGGCGA CAATGGTTAA AAGAAAACAG AGACCCATAT GTAAATTCTT AAGTCTGAGC

                                        BamHI
                                        ~~~~~~
3601  AGCAAGTCTA GAAAGGCGCG CCAAGATATC AAGGATCCAC TACGCGTTAG AGCTCGCTGA TCAGCCTCGA CTGTGCCTTC
      TCGTTCAGAT CTTTCCGCGC GGTTCTATAG TTCCTAGGTG ATGCGCAATC TCGAGCGACT AGTCGGAGCT GACACGGAAG

3681  TAGTTGCCAG CCATCTGTTG TTTGCCCCTC CCCCGTGCCT TCCTTGACCC TGGAAGGTGC CACTCCCACT GTCCTTTCCT
      ATCAACGGTC GGTAGACAAC AAACGGGGAG GGGGCACGGA AGGAACTGGG ACCTTCCACG GTGAGGGTGA CAGGAAAGGA

3761  AATAAAATGA GGAAATTGCA TCGCATTGTC TGAGTAGGTG TCATTCTATT CTGGGGGGTG GGGTGGGGCA GGACAGCAAG
      TTATTTTACT CCTTTAACGT AGCGTAACAG ACTCATCCAC AGTAAGATAA GACCCCCCAC CCCACCCCGT CCTGTCGTTC
```

FIG. 4D

```
3841  GGGGAGGATT GGGAAGACAA TAGCAGGCAT GCTGGGGAGC TCTTCCGCTT CCTCGCTCAC TGACTCGCTG CGCTCGGTCG
      CCCCTCCTAA CCCTTCTGTT ATCGTCCGTA CGACCCCTCG AGAAGGCGAA GGAGCGAGTG ACTGAGCGAC GCGAGCCAGC

3921  TTCGGCTGCG GCGAGCGGTA TCAGCTCACT CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG
      AAGCCGACGC CGCTCGCCAT AGTCGAGTGA GTTTCCGCCA TTATGCCAAT AGGTGTCTTA GTCCCCTATT GCGTCCTTTC

4001  AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC GTTGCTGGCG TTTTTCCATA GGCTCCGCCC
      TTGTACACTC GTTTTCCGGT CGTTTTCCGG TCCTTGGCAT TTTTCCGGCG CAACGACCGC AAAAAGGTAT CCGAGGCGGG

4081  CCCTGACGAG CATCACAAAA ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC CAGGCGTTTC
      GGGACTGCTC GTAGTGTTTT TAGCTGCGAG TTCAGTCTCC ACCGCTTTGG GCTGTCCTGA TATTTCTATG GTCCGCAAAG

4161  CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA
      GGGGACCTTC GAGGGAGCAC GCGAGAGGAC AAGGCTGGGA CGGCGAATGG CCTATGGACA GGCGGAAAGA GGGAAGCCCT

4241  AGCGTGGCGC TTTCTCAATG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC TCCAAGCTGG GCTGTGTGCA
      TCGCACCGCG AAAGAGTTAC GAGTGCGACA TCCATAGAGT CAAGCCACAT CCAGCAAGCG AGGTTCGACC CGACACACGT

4321  CGAACCCCCC GTTCAGCCCG ACCGCTGCGC CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT
      GCTTGGGGGG CAAGTCGGGC TGGCGACGCG GAATAGGCCA TTGATAGCAG AACTCAGGTT GGGCCATTCT GTGCTGAATA

4401  CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG
      GCGGTGACCG TCGTCGGTGA CCATTGTCCT AATCGTCTCG CTCCATACAT CCGCCACGAT GTCTCAAGAA CTTCACCACC

4481  CCTAACTACG GCTACACTAG AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA AAAGAGTTGG
      GGATTGATGC CGATGTGATC TTCCTGTCAT AAACCATAGA CGCGAGACGA CTTCGGTCAA TGGAAGCCTT TTTCTCAACC

4561  TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA
      ATCGAGAACT AGGCCGTTTG TTTGGTGGCG ACCATCGCCA CCAAAAAAAC AAACGTTCGT CGTCTAATGC GCGTCTTTTT

4641  AAGGATCTCA AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA ACTCACGTTA AGGGATTTTG
      TTCCTAGAGT TCTTCTAGGA AACTAGAAAA GATGCCCCAG ACTGCGAGTC ACCTTGCTTT TGAGTGCAAT TCCCTAAAAC

4721  GTCATGAGAT TATCAAAAAG GATCTTCACC TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA
      CAGTACTCTA ATAGTTTTTC CTAGAAGTGG ATCTAGGAAA ATTTAATTTT TACTTCAAAA TTTAGTTAGA TTTCATATAT

4801  TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA TCTCAGCGAT CTGTCTATTT CGTTCATCCA
      ACTCATTTGA ACCAGACTGT CAATGGTTAC GAATTAGTCA CTCCGTGGAT AGAGTCGCTA GACAGATAAA GCAAGTAGGT

4881  TAGTTGCCTG ACTCCCCGTC GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC AATGATACCG
      ATCAACGGAC TGAGGGGCAG CACATCTATT GATGCTATGC CCTCCCGAAT GGTAGACCGG GGTCACGACG TTACTATGGC

4961  CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC
      GCTCTGGGTG CGAGTGGCCG AGGTCTAAAT AGTCGTTATT TGGTCGGTCG GCCTTCCCGG CTCGCGTCTT CACCAGGACG

5041  AACTTTATCC GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC GCCAGTTAAT AGTTTGCGCA
      TTGAAATAGG CGGAGGTAGG TCAGATAATT AACAACGGCC CTTCGATCTC ATTCATCAAG CGGTCAATTA TCAAACGCGT

5121  ACGTTGTTGC CATTGCTACA GGCATCGTGG TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA
      TGCAACAACG GTAACGATGT CCGTAGCACC ACAGTGCGAG CAGCAAACCA TACCGAAGTA AGTCGAGGCC AAGGGTTGCT

5201  TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC CTTCGGTCCT CCGATCGTTG TCAGAAGTAA
      AGTTCCGCTC AATGTACTAG GGGGTACAAC ACGTTTTTTC GCCAATCGAG GAAGCCAGGA GGCTAGCAAC AGTCTTCATT

5281  GTTGGCCGCA GTGTTATCAC TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA AGATGCTTTT
      CAACCGGCGT CACAATAGTG AGTACCAATA CCGTCGTGAC GTATTAAGAG AATGACAGTA CGGTAGGCAT TCTACGAAAA
```

FIG. 4E

```
5361  CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAATA
      GACACTGACC ACTCATGAGT TGGTTCAGTA AGACTCTTAT CACATACGCC GCTGGCTCAA CGAGAACGGG CCGCAGTTAT

5441  CGGGATAATA CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT TCGGGGCGAA AACTCTCAAG
      GCCCTATTAT GGCGCGGTGT ATCGTCTTGA AATTTTCACG AGTAGTAACC TTTTGCAAGA AGCCCCGCTT TTGAGAGTTC

5521  GATCTTACCG CTGTTGAGAT CCAGTTCGAT GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA
      CTAGAATGGC GACAACTCTA GGTCAAGCTA CATTGGGTGA GCACGTGGGT TGACTAGAAG TCGTAGAAAA TGAAAGTGGT

5601  GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA ATAAGGGCGA CACGGAAATG TTGAATACTC
      CGCAAAGACC CACTCGTTTT TGTCCTTCCG TTTTACGGCG TTTTTTCCCT TATTCCCGCT GTGCCTTTAC AACTTATGAG

5681  ATACTCTTCC TTTTTCAATA TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
      TATGAGAAGG AAAAAGTTAT AATAACTTCG TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT

5761  GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA
      CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG GACTGCAGAT TCTTTGGTAA TAATAGTACT

5841  CATTAACCTA TAAAAATAGG CGTATCACGA GGCCCTTTCG TC
      GTAATTGGAT ATTTTTATCC GCATAGTGCT CCGGGAAAGC AG
```

FIG. 4F

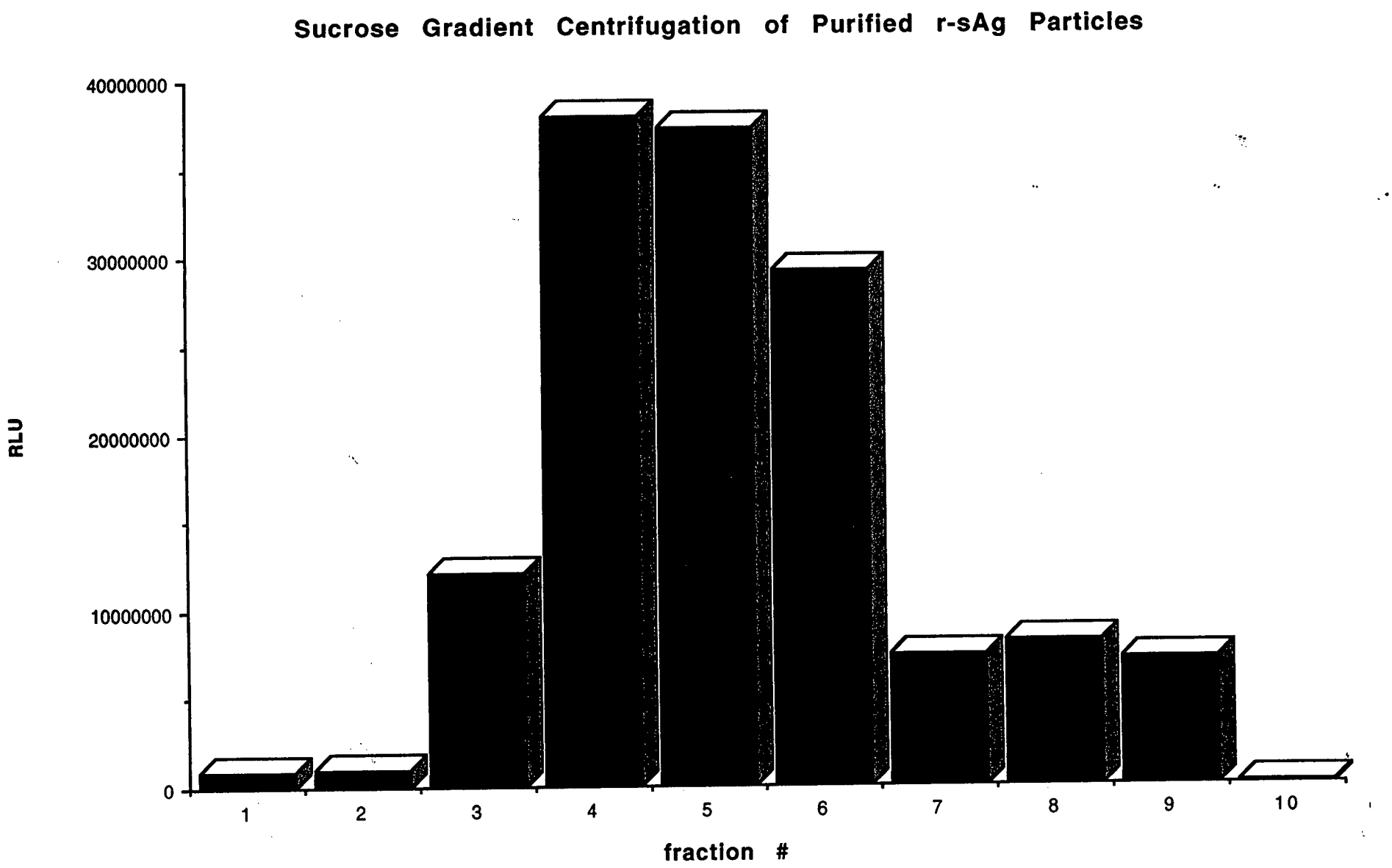
Sucrose Gradient Centrifugation of Purified r-sAg Particles
40000000
30000000
20000000
10000000
0
RLU
1
2
3
4
5
6
7
8
9
10
fraction #

Figure 5

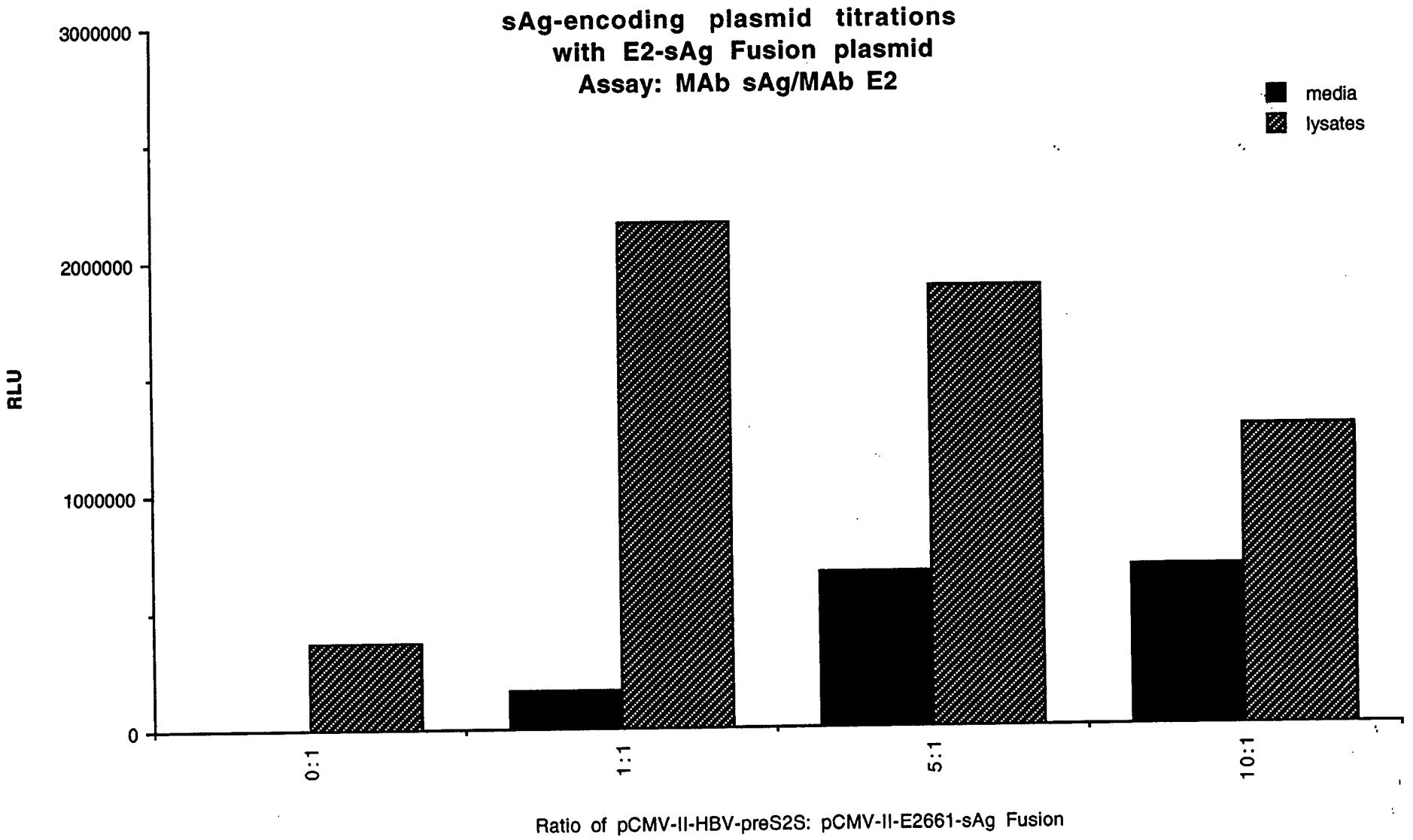
sAg-encoding plasmid titrations
with E2-sAg Fusion plasmid
Assay: MAb sAg/MAb E2
media
lysates
3000000
2000000
1000000
0
RLU
0:1
1:1
5:1
10:1
Ratio of pCMV-II-HBV-preS2S: pCMV-II-E2661-sAg Fusion

Figure 6A

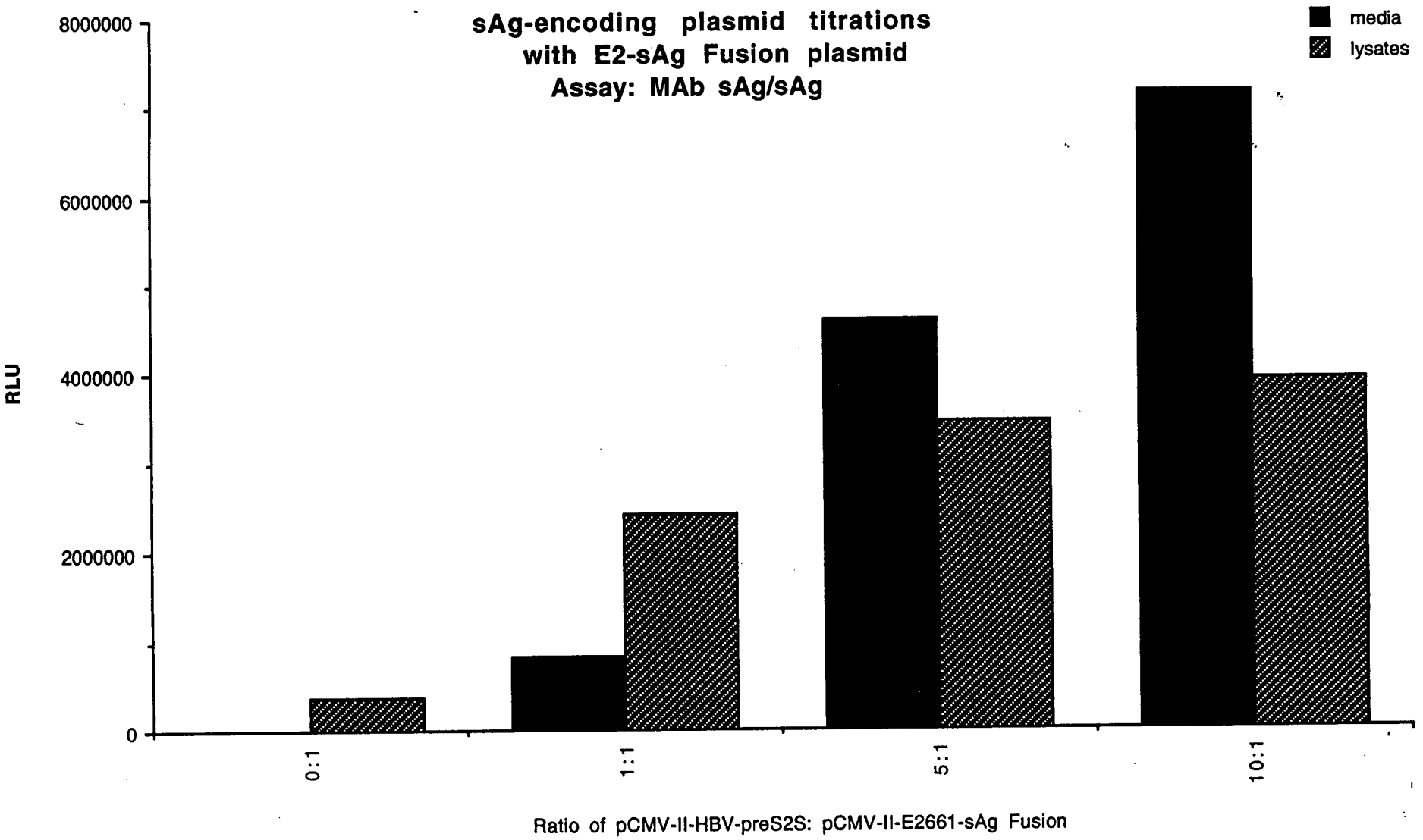
sAg-encoding plasmid titrations
with E2-sAg Fusion plasmid
Assay: MAb sAg/sAg
media
lysates
8000000
6000000
4000000
2000000
0
RLU
0:1
1:1
5:1
10:1
Ratio of pCMV-II-HBV-preS2S: pCMV-II-E2661-sAg Fusion

Figure 6B

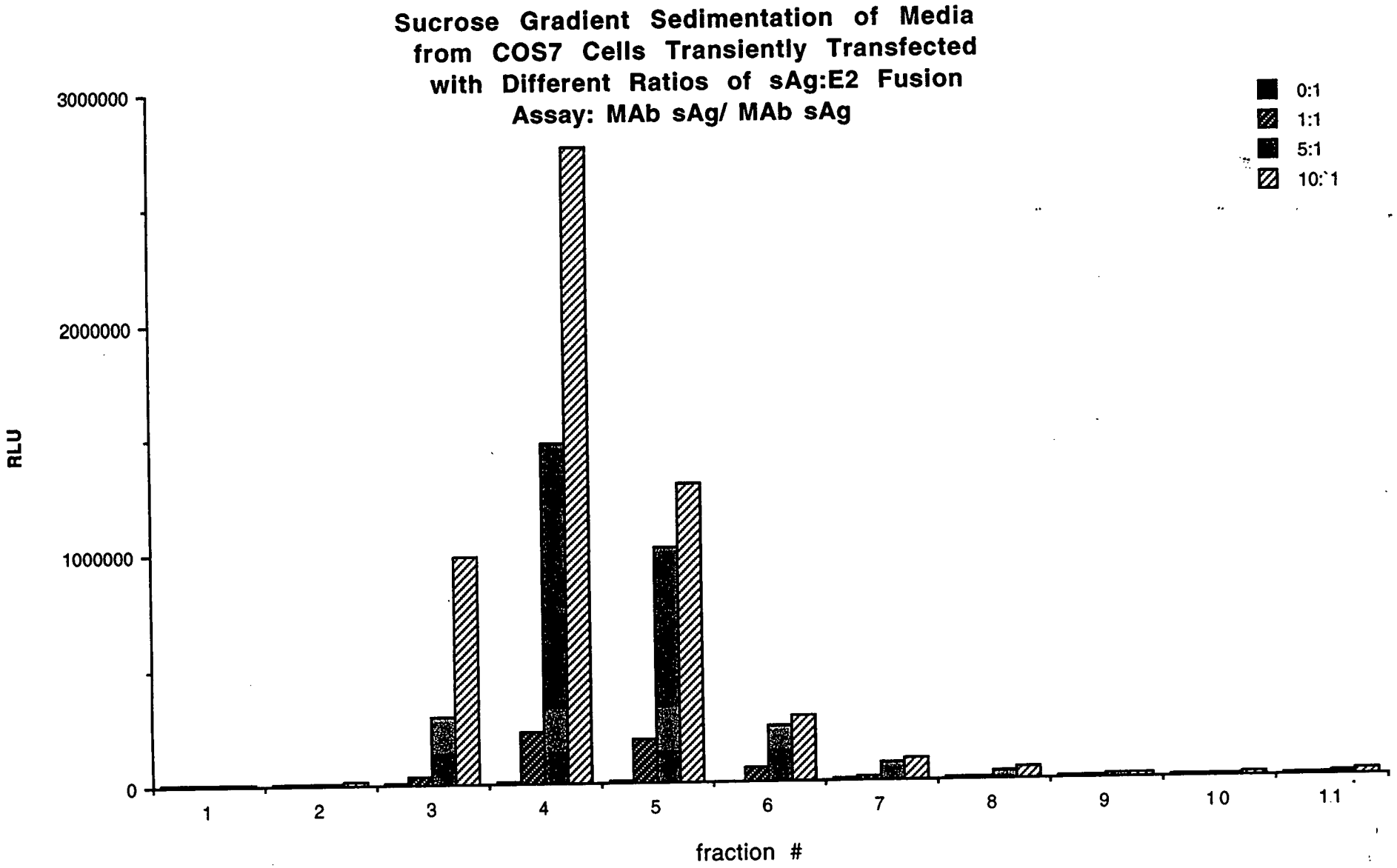
Sucrose Gradient Sedimentation of Media from COS7 Cells Transiently Transfected with Different Ratios of sAg:E2 Fusion Assay: MAb sAg/ MAb sAg
0:1
1:1
5:1
10:1
3000000
2000000
1000000
0
RLU
1
2
3
4
5
6
7
8
9
10
11
fraction #

Figure 7A

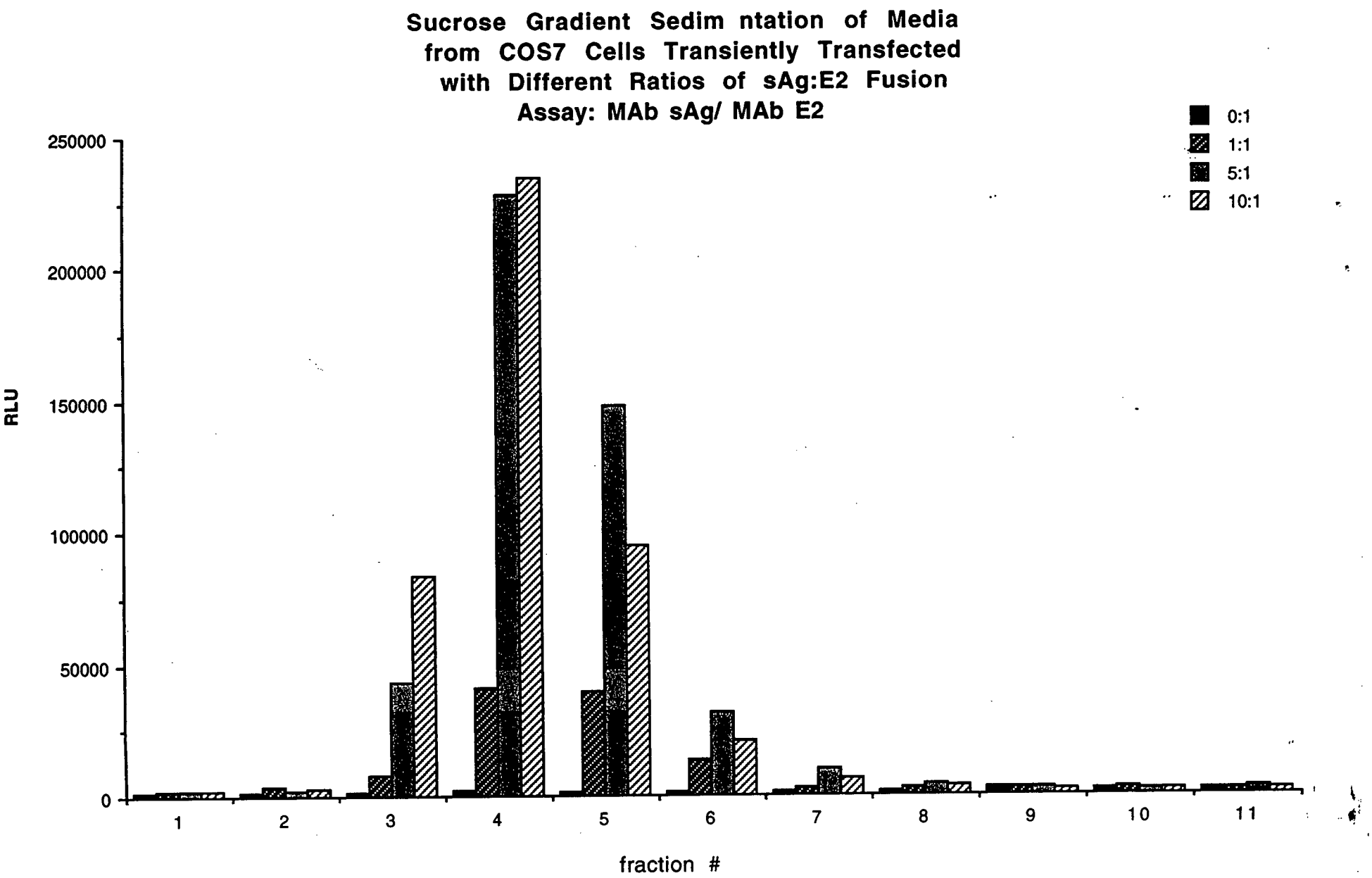
Sucrose Gradient Sedim ntation of Media from COS7 Cells Transiently Transfected with Different Ratios of sAg:E2 Fusion Assay: MAb sAg/ MAb E2
0:1
1:1
5:1
10:1
RLU
250000
200000
150000
100000
50000
0
1
2
3
4
5
6
7
8
9
10
11
fraction #

Figure 7B

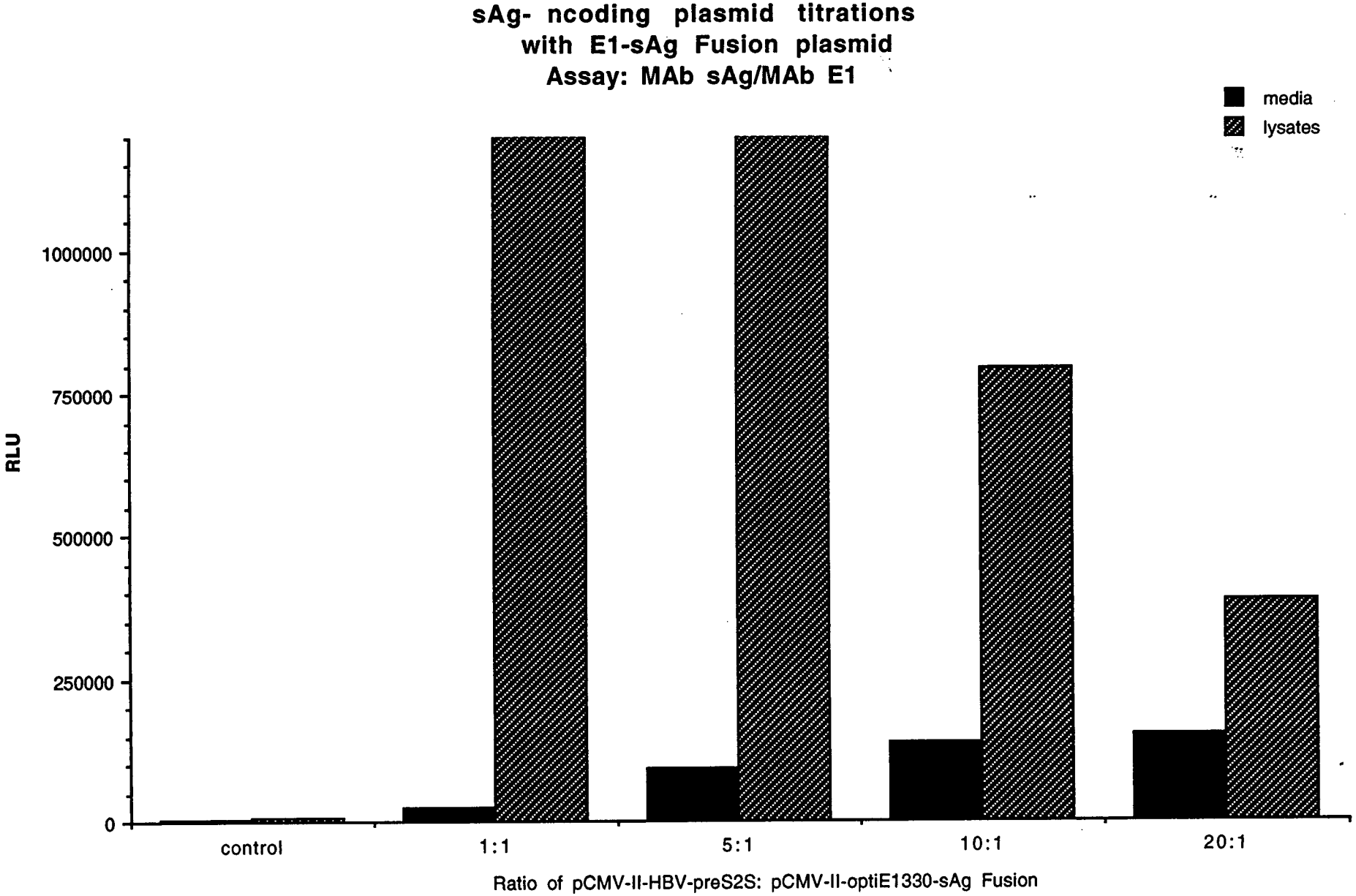
sAg- ncoding plasmid titrations
with E1-sAg Fusion plasmid
Assay: MAb sAg/MAb E1
media
lysates
RLU
1000000
750000
500000
250000
0
control
1:1
5:1
10:1
20:1
Ratio of pCMV-II-HBV-preS2S: pCMV-II-optiE1330-sAg Fusion

Figure 8A

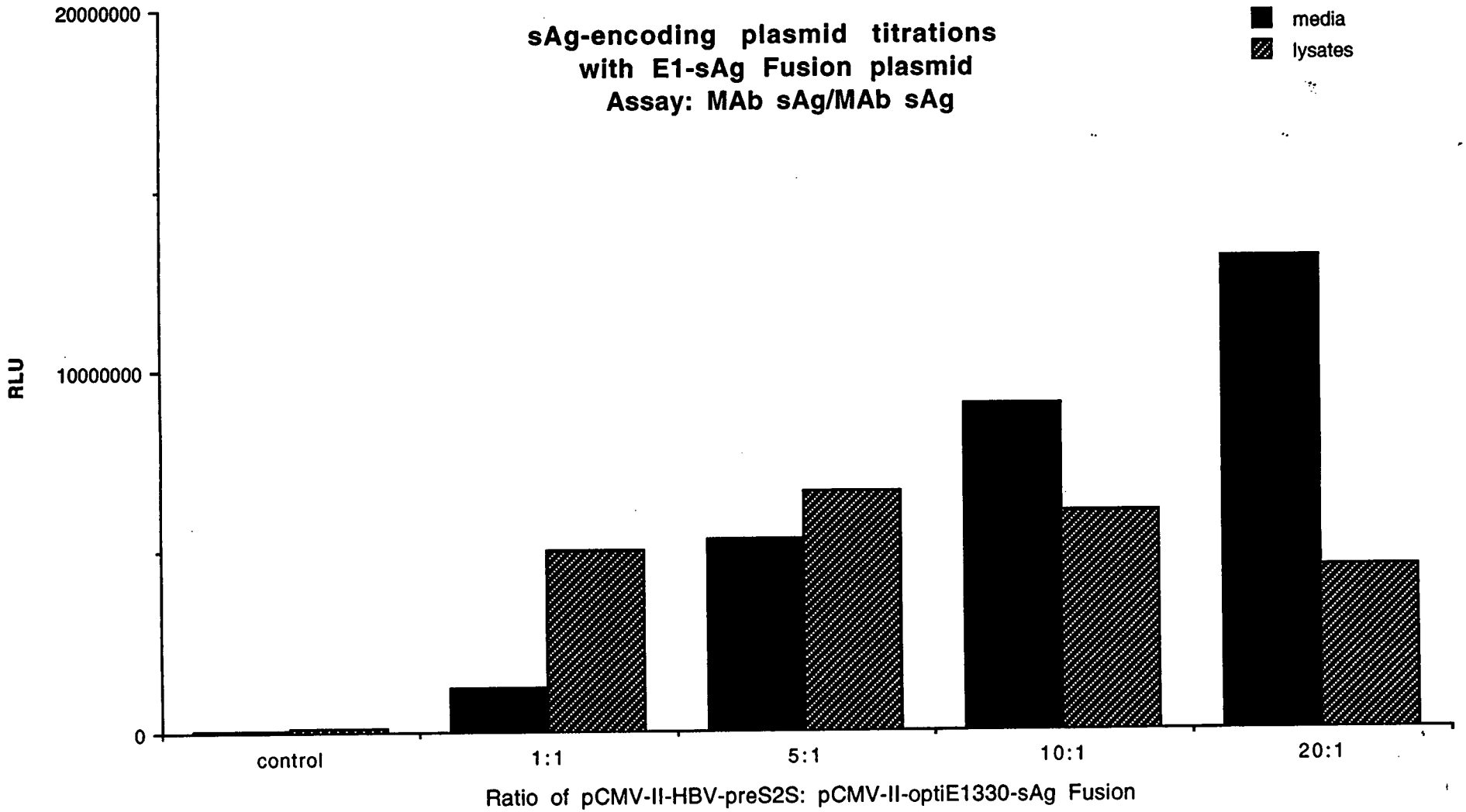

Figure 8B

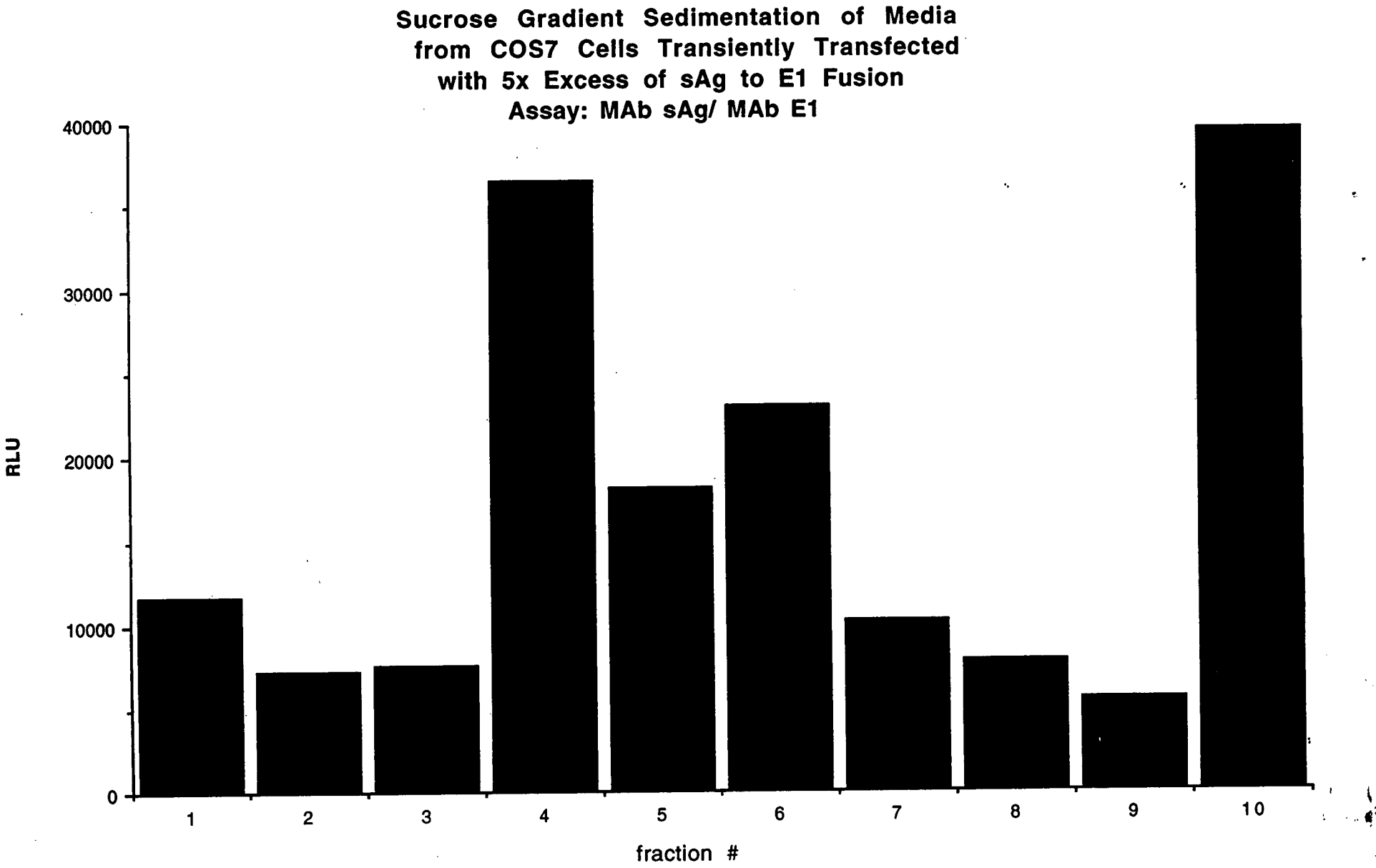

Figure 9A

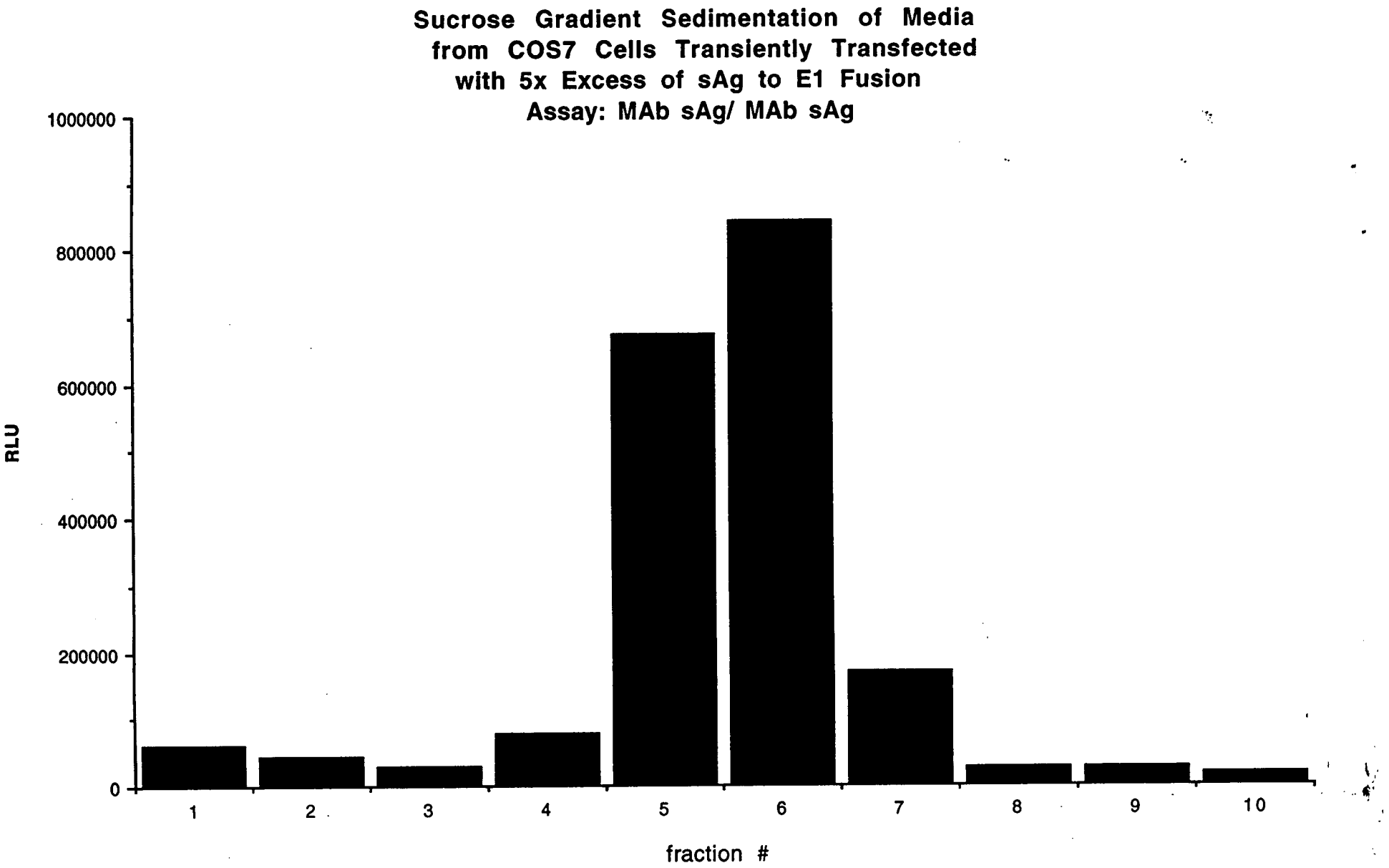

Figure 9B

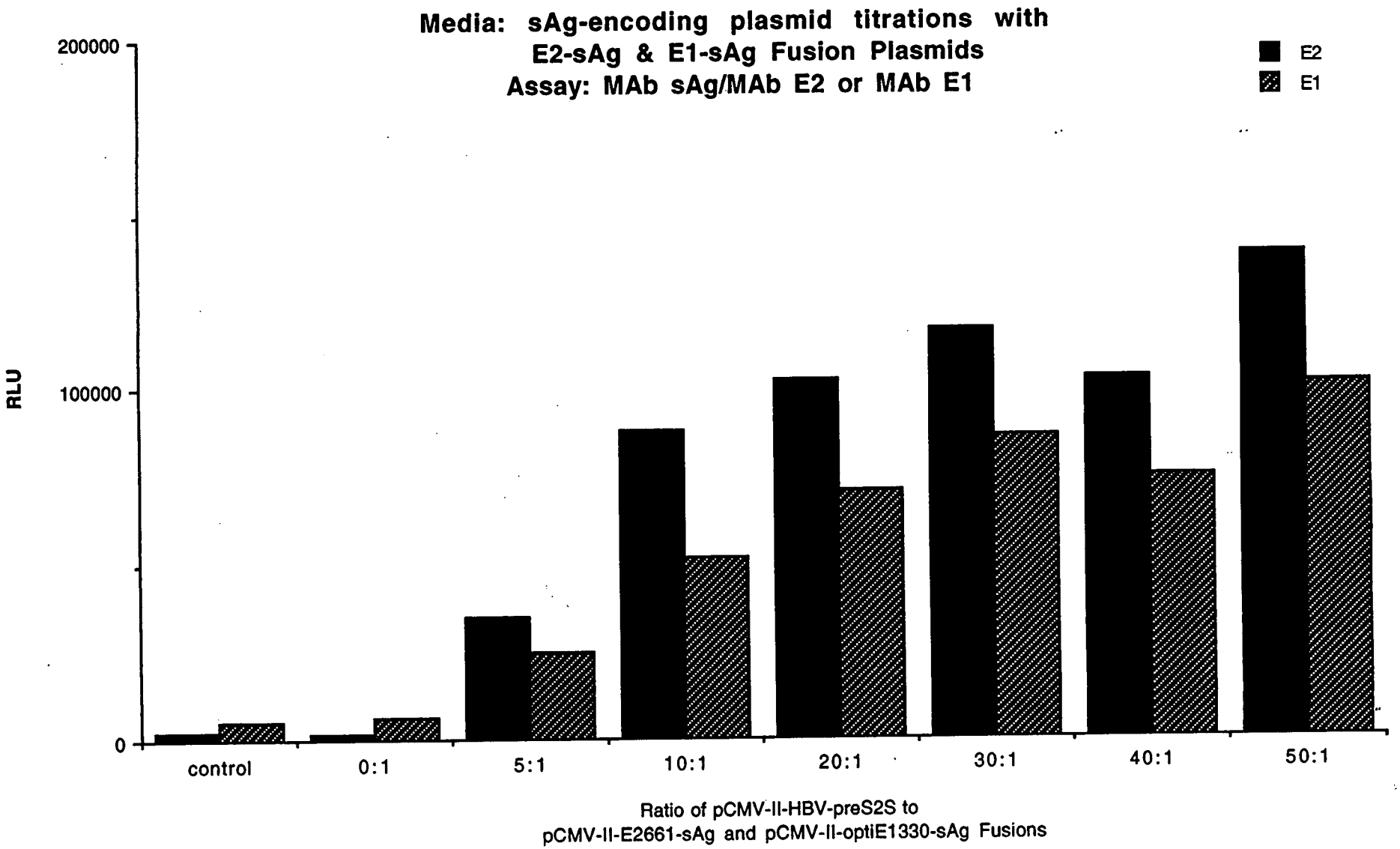

Figure 10A

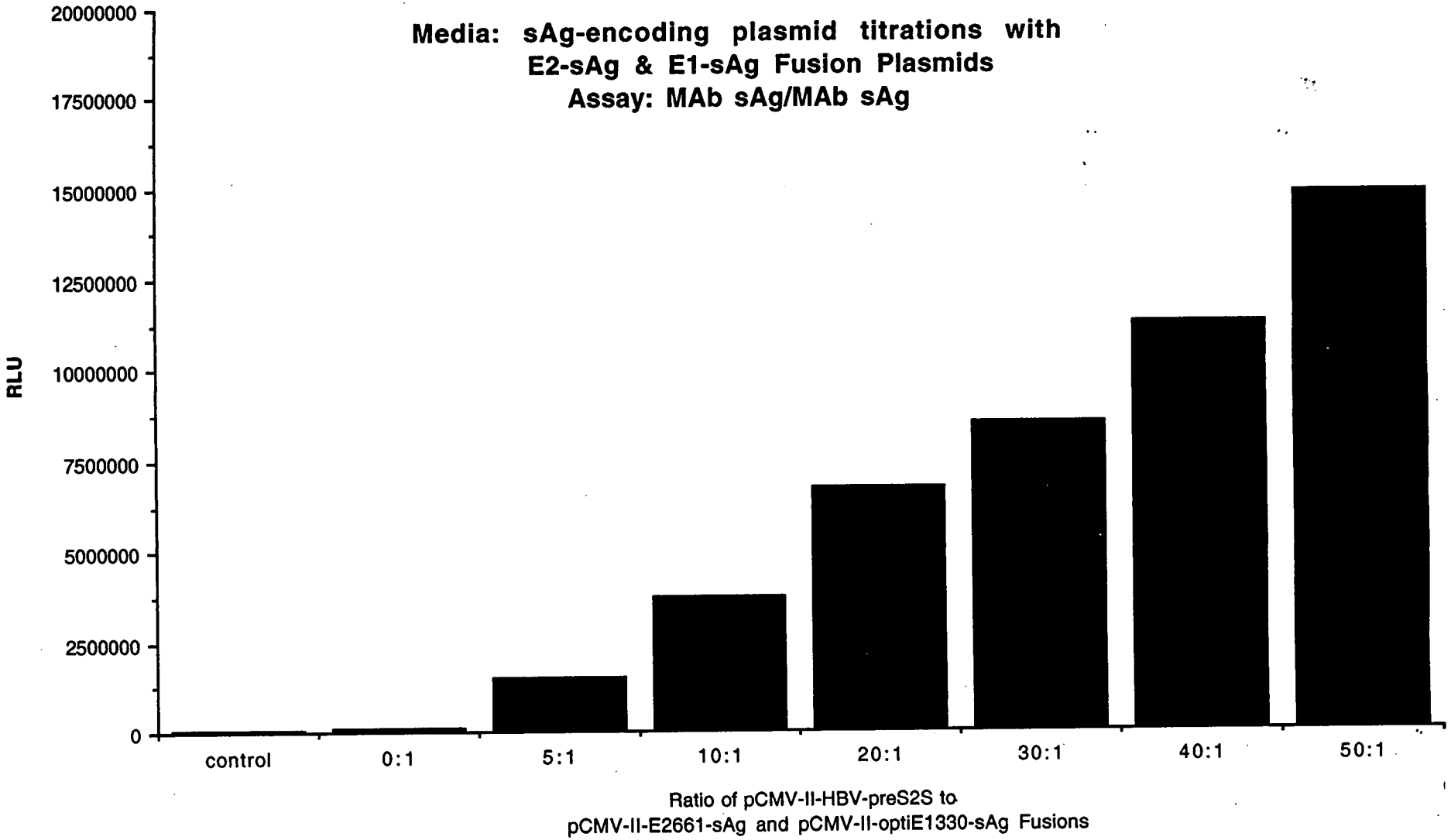

Figure 10B